2012

Equity performance evaluation of two different pricing options: fuel tax per gallon and VMT fee

Eirini Kastrouni
Iowa State University

Equity performance evaluation of two different pricing options:

fuel tax per gallon and VMT fee

by

Eirini Kastrouni

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee:
Konstantina Gkritza, Major Professor
Shauna Hallmark
Robert W. Stephenson

Iowa State University

Ames, Iowa

2012

Dedication

I would like to thank my wonderful parents, Ilias and Vassiliki, and my brother, Michalis. Their love, support, and encouragement have been essential for me to achieve my professional and personal goals. I would also like to thank A. for always being there for me.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank Dr. Konstantina Gkritza, who has served as my Major Professor during my graduate studies. Her support and contribution have been essential in developing the research presented in this thesis, and she has been a mentor for me throughout these two years in all levels of academic life.

I would also like to thank my Program of Study Committee Members, Dr. Shauna Hallmark and Dr. Robert W. Stephenson for their useful comments and suggestions which shaped my final research analysis.

My gratitude is extended to Dr. James McCalley for the financial aid provided to support my graduate studies, through NETSCORE21.

I would also like to thank Inya Nlenanya for his assistance with data processing.

Finally, I would like to thank Dr. Matthew Karlaftis, my mentor in Greece; his encouragement and support have been valuable all along my academic life.

Abstract

Currently, there has been extensive discussion on various pricing alternatives to the present tax per fuel gallon that is in effect. There has been great interest to identify additional resources that will increase the federal and state revenues allocated to the maintenance of the existing surface transportation network. The widely suggested policy measure of the vehicle-miles-traveled fee (VMT fee) is an alternative pricing option that has drawn great attention by researchers and policymakers, particularly regarding its equity performance among various social groups.

In this context, the objective of this thesis is two-fold. The primary objective is to identify which social sub-groups are mostly affected under the current fuel tax option and the alternative VMT fee option. To achieve this, the author collated information on socioeconomic-, geographic-, and vehicle-specific attributes at the household (HH) level from the original Household and Vehicle Files of the 2009 National Household Travel Survey. The identification of the social sub-groups is realized via a three-stage least squares (3SLS) model development at the national level of analysis, where the dependent variables in the model specification for each pricing option are the average vehicle fuel efficiency and the vehicle-miles traveled for the fuel tax and the VMT fee option respectively.

The second research objective of this thesis is to identify if the model specification at the national level may be applicable at a more localized level of analysis, for example for Iowa. The analysis of variance (ANOVA) and the asymptotic t-test of the individual coefficients are applied to test for differences in the two levels of analysis.

The results of the first part of the analysis show that particular social sub-groups, such as HHs located in rural areas, or HHs that are located in states with lower fuel taxation,

operate vehicles of lower fuel efficiency at the HH level, thus consume more fuel for the same distance traveled, and therefore have higher fuel-related expenditures. Such groups are expected to shoulder greater of the fuel tax burden, as their relation to the average vehicle fuel efficiency at the HH level via the model specification is decreasing. On the other hand, based on the second national model specification, HHs such as those that own vehicles of higher fuel efficiency, or are located in rural areas, or have a higher average income generate more trips at an annual basis, thus have a higher VMT at the HH level.

Regarding the second part of the analysis, the differences between the national and local model lie in three different levels, suggesting that, despite the similarities, the development of a distinct local model is statistically supported.

- There are statistically significant differences in the VMT at the HH level, suggesting that VMT in Iowa exhibit different trend than the national average VMT.
- The two levels of analysis share quite a few common variables in the model specification, but the location-specific variables did not participate in the local model development.
- Based on the asymptotic t-test of equality of individual coefficients, there is a statistically significant difference detected in the coefficient estimates of all variables, suggesting that their effect magnitude on the average vehicle fuel efficiency and VMT at the HH level differs between the national and the local model.

Keywords: Fuel tax per gallon, VMT fee, equity, three-stage least squares, 2009 NHTS.

Chapter 1: Introduction

1.1. Research Motivation

Lately, there has been a lot of discussion with respect to the limited revenues that are available to be allocated to the surface transportation system and its maintenance. The Highway Trust Fund (HTF) is not sufficient to cover the financial needs of the surface transportation system (TRB Committee for the Study of the Long-Term Viability of Fuel Taxes for Transportation Finance, 2006), (TRB Committee on Equity Implications of Evolving Transportation Finance Mechanisms, 2011), (National Surface Transportation Policy and Revenue Study Commission, 2007), (National Surface Transportation Infrastructure Financing Commission, 2009). Initiated in 1956, the current federal surface transportation bill cannot successfully meet the needs of the modern transportation system, given that there is limited political will to increase the federal gas tax. SAFETEA-LU (Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users) is the latest authorization bill that was signed into law in August 10th, 2005 and is currently under its 9th extension, till May 2012, since its expiration back in September 2009. In September 2008, the Highway Trust Fund was backfilled with $8 billion in general revenues so as to maintain the transportation program. The Congress proceeded with the development of the National Surface Transportation Infrastructure Financing Commission; the generated report suggests a $0.10 increase in fuel tax to render HTF solvent and transition to a mileage-based fee. Still, however, Shank & Rudnick-Thorpe (2011) consider this pricing option to be unsustainable in its performance, and put forward the need to proceed with practices which meet the economic, environmental and social criteria of sustainability.

Within the context of insufficiency of the current sources of revenues, there has been a lot of discussion regarding other potential sources of revenue, proceeding from different revenue-related policy measures, such as the VMT fee, feebates, and rebates. Each pricing option may have a great number of different implementation schemes, depending on the setting of its parameters (level of fee, implementation of policy for particular groups, or other). The greatest part of the discussion has focused on two distinct pricing options: increasing the current fuel tax, and implementing the suggested vehicle-miles traveled (VMT) fee. However, as with most newly suggested policy measures, there has been some conflict regarding the impacts that each alternative may have on different groups, and whether these alternatives are equitable (or socially fair).

This discussion, which will be further presented in Chapter 2, has motivated the author in addressing the equity impacts of each alternative, within the broader context of transportation sustainability. Defining transportation sustainability is challenging due to the inherently broad nature of the concept. The traditional definition of sustainability calls for policies and strategies "that meet society's present needs without compromising the ability of future generations to meet their own needs" (U.S. Environmental Protection Agency). In this thesis, the author uses a second widely accepted definition of sustainable development as has been provided by the Brundtland Commission. According to this definition, "*sustainable development seeks to meet the needs and aspirations of the present without compromising the ability to meet those of the future"* (United Nations, 1987)*)*. This implies *"the balancing of economic, social, environmental and technological considerations as well as the incorporation of a set of ethical values"* (Council of Academies of Engineering and Technological Sciences, 1995). Thus, as illustrated in Figure 1-1, sustainability contains three overlapping dimensions: environmental, social, and economic (The World Conservation Union, 2006), (International Atomic Energy Agency, 2005), (Vera, Langlois, Rogner, Jalal, & Toth, 2005), (Whang, Jinga, Zhanga, & Zhaoa, 2009).

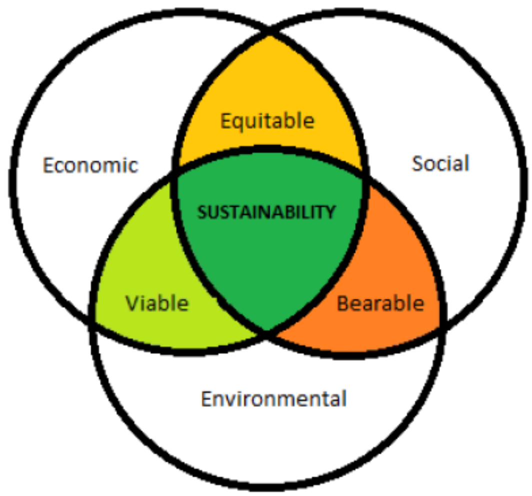

Figure 1-1: The three dimensions of sustainability (Adapted from (United Nations, 1987)

The concept of sustainability is comprised of three pillars: environment, economy, and society. A sustainable practice or a sustainable project is the one that meets the sustainability needs in all three fields. A practice that performs sustainably in terms of its environmental and economic impacts should not be considered sustainable in total, if it ignores the third pillar of sustainability, i.e. society.

The two tax options which are studied herein are assessed in terms of their performing in accordance with the principles of the third pillar of sustainability, i.e. the author attempts to assess their impact on particular social groups.

Although the environmental and economic aspects will not be addressed in this thesis, it is noted that the environmental aspect of sustainability is the one that has received the largest amount of attention by researchers, professionals, and policymakers. Falsely, it is treated as the single goal of sustainable performance, although theoretically and practically all three aspects of sustainability should receive equal attention and should be managed in parallel. The effect of the two tax alternatives on environment is expected to vary, since there are inherent differences between those two. The current practice of fuel tax per gallon is closely associated to the vehicle type and its fuel efficiency. Fuel efficient vehicles receive lower taxation per gallon consumed, compared to less efficient vehicles, since their technology allows for traveling a certain distance by consuming fewer gallons of fuel. Such an advantage is thought to induce more travel, more VMT, and thus a stronger and negative impact on the environment. On the other hand, a VMT related fee clearly sets some limitations; road users will be charged according to the extent they use the road infrastructure, discouraging them from generating unnecessary trips. Such a policy measure is expected to impact the overall VMT, in a more straightforward way, leading to lower vehicle-related emissions.

Similarly, the long-term economic effect of each tax alternative can be assessed in the sense of the revenue it is expected to generate. According to previous work, the fuel tax option seems inadequate in generating the necessary revenue to support the road infrastructure system. On the other hand, the mileage fee is a widely recommended tax alternative which endeavors to generate the necessary revenue and fill the gap that past tax options created.

The social aspect of sustainability is usually the hardest to capture, and is the main focus of this thesis. Usually, the social effects of a policy measure are hard to quantify, hindering policymakers from assessing the measure's real impact on society. In this analysis, the social impact of each alternative is quantified by identifying which social groups are mostly affected under each pricing option. For example, in the current case of fuel tax, the social groups which are expected to be affected are those that do not have easy access to fuel efficient vehicles. Those groups may be race-, income-, or age-specific, suggesting that different demographic and socioeconomic will experience dissimilar conditions. Similarly, in the case of a mileage fee, the groups that may be affected may be those who reside in rural areas (leading to higher daily VMT) and do not have access to public transit (e.g. rail etc.)

1.2. Thesis objectives

The main objective of this study is to assess the performance of these two different pricing alternatives with respect to the concept of sustainability. Focusing on the social aspect of sustainability, the author is interested in identifying which social subgroups are more likely to be affected under each pricing option. Making inferences and/or suggestions regarding which alternative is best for state agencies or the federal government to use is not the main objective of this study; however, evaluating each option's equity impacts may prove valuable to policymakers and other stakeholders who wish to ensure that the policy measures implemented by their agencies are equitable.

A secondary objective of this thesis is to obtain model estimations for both a national and a local model (Iowa), in a bid to identify if there are inherent differences in the variables of interest between the two levels of analysis, and whether a single model specification may successfully describe reality and yield valid predictions for the future, at both levels of analysis. The model specification at both scales of analysis shall include socio-economic-, vehicle-, and geographic-specific variables, in order to capture various aspects of the HH characteristics.

1.3. Thesis structure

This thesis is organized into five chapters.

Chapter 2: *Literature Review* includes an overview of previous studies relating to various pricing options. This chapter mainly focuses on the fuel tax and the VMT fee alternatives; however the author provides an overview of studies which have focused on other pricing alternatives as well, such as feebates, rebates, and other.

Chapter 3: *Data Description* provides details about the two datasets (national and local for Iowa) used in order to conduct the proposed analysis. The means to describe the data include descriptive statistics, correlation matrices and plots of the variables of interest, to help visualize each one of the two pricing options under evaluation.

Chapter 4: *Methodology and Results* discusses the methodology applied for the model development, as well as for the level-of-analysis comparison. Additionally, this chapter presents the results of the conducted analysis, for each pricing option both at the national and local level, as well as the results of the model comparison between the national and the local level of analysis.

Finally, Chapter 5: *Conclusions, Limitations and Recommendations* offers concluding remarks on the analysis, as well as information on the limitations of this study and suggestions for future research.

Chapter 2: Literature Review

2.1. Overview

This chapter discusses various pricing options that have drawn the attention of both researchers and policymakers, in a bid to identify the benefits and constraints of their performance, the requirements for their implementation, and their effects on society, in a broader context. The current debate mainly pertains to two pricing options for long-term transportation funding at either the federal or the state level; fuel tax per gallon, and the VMT fee. However, this chapter also discusses in brief some other pricing options that have received some attention, in an attempt to provide a more comprehensive review of the available pricing options.

2.2. Fuel tax per gallon

2.2.1. Overview

The current state-of-practice regarding the road-infrastructure-related taxation is the "fuel tax per gallon". According to this taxation practice, users are charged at-the-pump with respect to how many gallons of fuel they purchase. The fuel rate consists of a federal excise tax at $0.184 per gallon, plus a state tax, which varies depending on the state. Table 2-1 provides information on the rates of the federal and state taxes by state. The column ''Other State Taxes/Fees'' includes different types taxes/fees, as described in detail in the report prepared by API (American Petroleum Institute). For example, in Iowa, "Other Taxes" column includes 1 cpg Underground Storage Tank (UST) fee.

Table 2-1: State Motor Fuel (Gasoline) Excise and Other Taxes (American Petroleum Institute)

State	State Excise Tax (cpg)	Other State Taxes/ Fees (cpg)	Total State Taxes/ Fees (cpg)	Total State plus Federal Excise Taxes (at 18.4 cpg)
Alabama	16.0	4.9	20.9	39.3
Alaska	8.0	0.0	8.0	26.4
Arizona	18.0	1.0	19.0	37.4
Arkansas	21.5	0.3	21.8	40.2
California	35.7	13.4	49.1	67.5
Colorado	22.0	0.0	22.0	40.4
Connecticut	25.0	24.6	49.6	68.0
Delaware	23.0	0.0	23.0	41.4
District of Columbia	23.5	0.0	23.5	41.9
Florida	4.0	30.5	34.5	52.9
Georgia	7.5	21.7	29.2	47.6
Hawaii	17.0	30.4	47.4	65.8
Idaho	25.0	0.0	25.0	43.4
Illinois	19.0	22.2	41.2	59.6
Indiana	18.0	21.7	39.7	58.1
Iowa	21.0	1.0	22.0	40.4
Kansas	24.0	1.0	25.0	43.4
Kentucky	26.4	1.4	27.8	46.2
Louisiana	20.0	0.0	20.0	38.4
Maine	30.0	1.5	31.5	49.9
Maryland	23.5	0.0	23.5	41.9
Massachusetts	21.0	2.5	23.5	41.9
Michigan	19.0	21.8	40.8	59.2
Minnesota	27.1	0.1	27.2	45.6
Mississippi	18.0	0.8	18.8	37.2
Missouri	17.0	0.3	17.3	35.7
Montana	27.0	0.8	27.8	46.2
Nebraska	26.3	0.9	27.2	45.6
Nevada	23.0	10.1	33.1	51.5
New Hampshire	18.0	1.6	19.6	38.0
New Jersey	10.5	4.0	14.5	32.9
New Mexico	17.0	1.9	18.9	37.3

State	**State Excise Tax (cpg)**	**Other State Taxes/ Fees (cpg)**	**Total State Taxes/ Fees (cpg)**	**Total State plus Federal Excise Taxes (at 18.4 cpg)**
New York	8.1	41.4	49.5	67.9
North Carolina	35.0	0.3	35.3	53.7
North Dakota	23.0	0.0	23.0	41.4
Ohio	28.0	0.0	28.0	46.4
Oklahoma	16.0	1.0	17.0	35.4
Oregon	30.0	1.0	31.0	49.4
Pennsylvania	12.0	20.3	32.3	50.7
Rhode Island	32.0	1.0	33.0	51.4
South Carolina	16.0	0.8	16.8	35.2
South Dakota	22.0	2.0	24.0	42.4
Tennessee	20.0	1.4	21.4	39.8
Texas	20.0	0.0	20.0	38.4
Utah	24.5	0.0	24.5	42.9
Vermont	19.0	7.6	26.6	45.0
Virginia	17.5	2.5	20.0	38.4
Washington	37.5	0.0	37.5	55.9
West Virginia	20.5	11.7	32.2	50.6
Wisconsin	30.9	2.0	32.9	51.3
Wyoming	13.0	1.0	14.0	32.4
US Average	***20.8***	***9.7***	***30.5***	***48.9***

With an average of $0.489 per gallon, 17 states are above the mean, with Connecticut ranking the highest with $0.68 per gallon, and 34 states are below the average, with Alaska being the state with the lowest tax rate at $0.264 per gallon of gasoline. Figure 2-1 and Figure 2-2 show the total tax rates per gallon of fuel purchased by state.

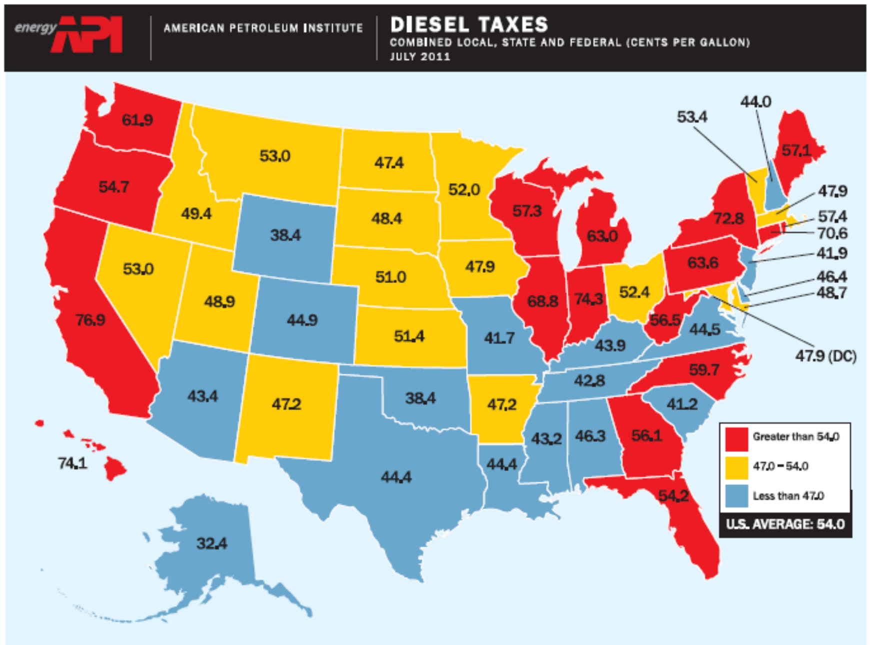

Figure 2-1: Diesel Taxes by State (American Petroleum Institute)

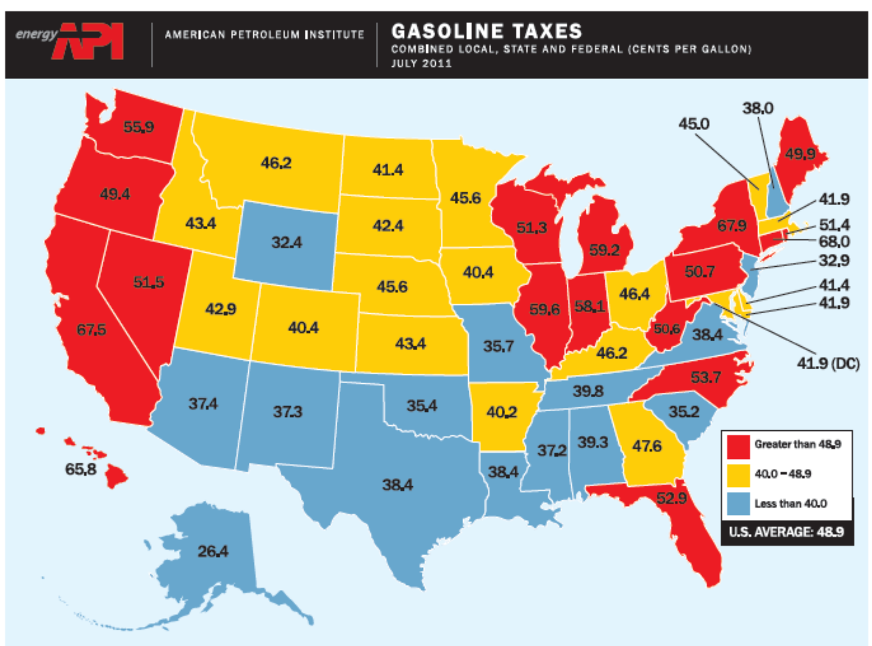

Figure 2-2: Gasoline Taxes by State (American Petroleum Institute)

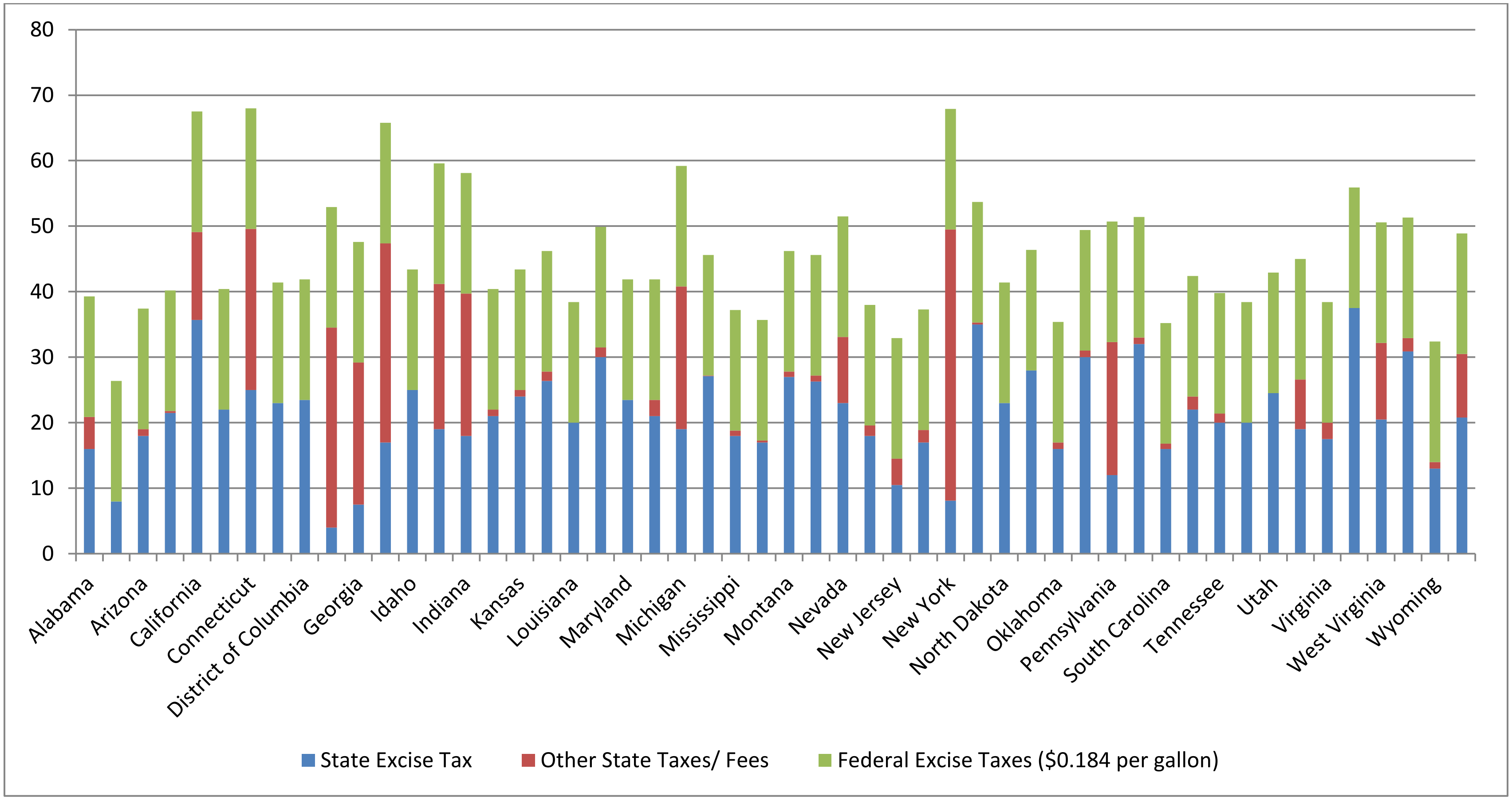

Figure 2-3: State Motor Fuel (Gasoline) Excise and Other Taxes
Data Source: (American Petroleum Institute)

The pricing option of fuel tax per gallon has started facing some opposition as it is not believed to be able to meet its targets with regards to revenue raised for the design, construction, operation, and maintenance of the surface transportation infrastructure. This pricing option leads to different user charges depending on the fuel efficiency of the vehicle. Vehicles with higher vehicle efficiency (in terms of miles per gallon – mpg) have lower fuel consumption, thus pay lower taxes compared to less fuel-efficient vehicles, for using the road infrastructure to the same extent (i.e. equal level of miles driven). Another issue that should also be noted is that the current federal tax rates of $0.184 per gallon of gasoline and $0.244 per gallon of diesel have not been increased or indexed to inflation since 1993, thus their ‘’purchasing power’’ has reduced even more during this long period.

2.2.2. CAFE standards

A configuration which has tremendous effects on the effectiveness of the fuel tax pricing option is the implementation of the CAFE (Corporate Average Fuel Economy) standards.

‘’*Corporate Average Fuel Economy (CAFE) is the sales weighted average fuel economy, expressed in miles per gallon (mpg), of a manufacturer's fleet of passenger cars or light trucks with a gross vehicle weight rating (GVWR) of 8,500 lbs. or less, manufactured for sale in the United States, for any given model year’’* (NHTSA). The primary objective of the CAFE standards is to increase the average fuel economy of cars and light trucks available in the market, thus reduce the associated energy consumption (NHTSA).

First enacted by Congress in 1975, those standards are established and amended by the National Highway Traffic Safety Administration (NHTSA) for every Model Year (MY) since 1978. The US Environmental Protection Agency (EPA) calculates the average fuel economy for each manufacturer. The fuel economy test data is either provided to EPA by the manufacturer, or obtained by EPA after testing the vehicle, in its Office of Transportation & Air Quality facility in Ann Arbor, MI. (NHTSA)

The most recent standards were issued in 2011 and there has been great amount of speculation in the press regarding future vehicle-fuel-efficiency goals, set by President Obama and other policymakers. President Obama's policy regarding the vehicle fuel efficiency standards is very aggressive; the target for 2025 is almost doubling the average fuel economy, from the current level of 30.2 to 54.5 mpg (Obama announces 54.5 mpg CAFE standard by 2025)

Historic data for the average fuel efficiency of the passenger car fleet from 1958 to 1978 is presented in Figure 2-4; the data infers that domestic manufacturers produced vehicles of lower fuel efficiency than their international counterparts, implying that political concern with regard to environment and energy conservation was belated in the U.S., in comparison with the rest of the world.

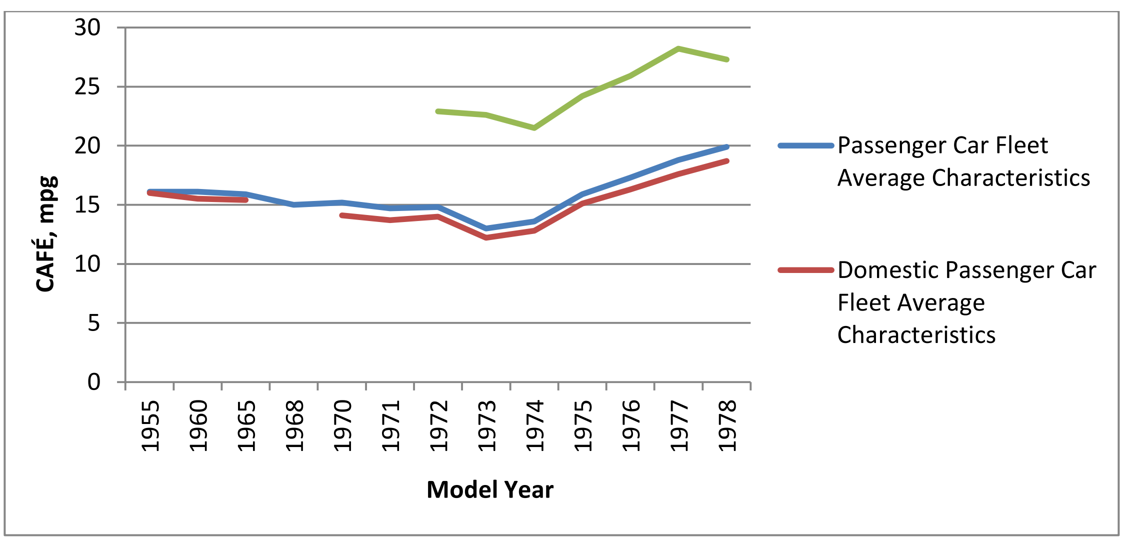

Figure 2-4: Historical Passenger Car Fleet Average Characteristics, 1955 – 1978 (NHTSA)

The low fuel efficiency of the passenger and light truck fleet pushed the politicians towards establishing and implementing the CAFE standards. Since 1978, distinct CAFE standards are set for passenger cars and light trucks. Those are listed in Table 2-2 for all Model Years, from 1978 to 2011.

Table 2-2: CAFE standards for passenger cars and light trucks (U.S. Department of Transportation, NHTSA, 2011)

Model Year	Passenger Cars	Light Trucks
		Combined 2WD & 4WD
1978	18.0	N/A
1979	19.0	N/A
1980	20.0	N/A
1981	22.0	N/A
1982	24.0	17.5
1983	26.0	19.0
1984	27.0	20.0
1985	27.5	19.5
1986	26.0	20.0
1987	26.0	20.5
1988	26.0	20.5
1989	26.5	20.5
1990	27.5	20.0
1991	27.5	20.2
1992	27.5	20.2
1993	27.5	20.4
1994	27.5	20.5
1995	27.5	20.6
1996	27.5	20.7
1997	27.5	20.7
1998	27.5	20.7
1999	27.5	20.7
2000	27.5	20.7
2001	27.5	20.7
2002	27.5	20.7
2003	27.5	20.7

Model Year	Passenger Cars	Light Trucks *Combined 2WD & 4WD*
2004	27.5	20.7
2005	27.5	21.0
2006	27.5	21.6
2007	27.5	22.2
2008	27.5	22.5
2009	27.5	23.1
2010	27.5	23.5
2011	30.2	24.2

Upon establishment of the standards, car-manufacturers, both domestic and international ones, had to comply with them, in a bid to avoid penalties and other restrictions. Figure 2-5 and Figure 2-6 suggest that since 1978, both passenger cars and light trucks' average fuel efficiency has exceeded the limit values, and perform better than required.

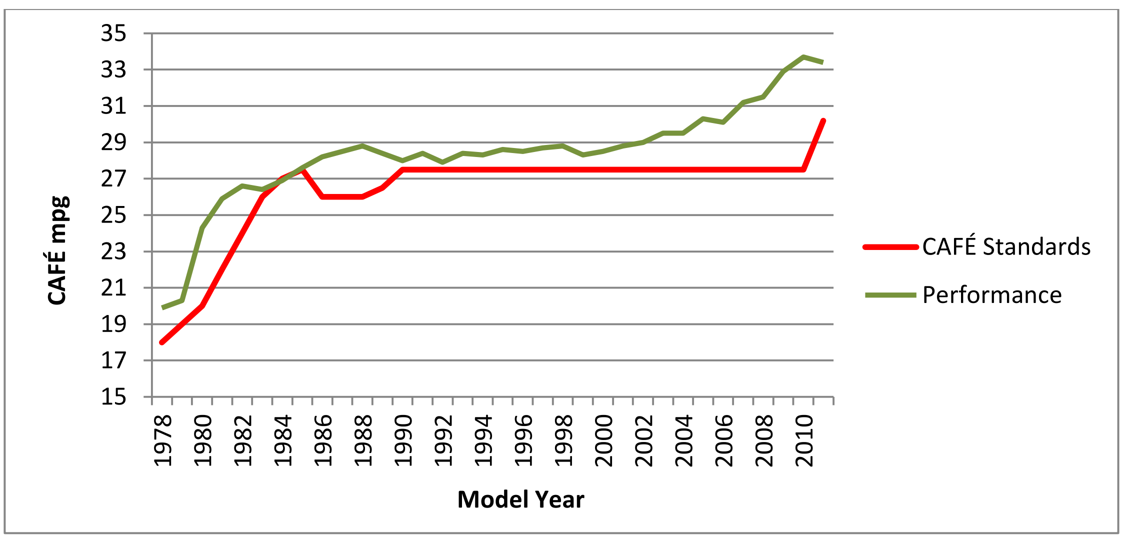

Figure 2-5: Passenger Car Fleet Performance, 1978-2011
Data Source: (U.S. Department of Transportation, NHTSA, 2011)

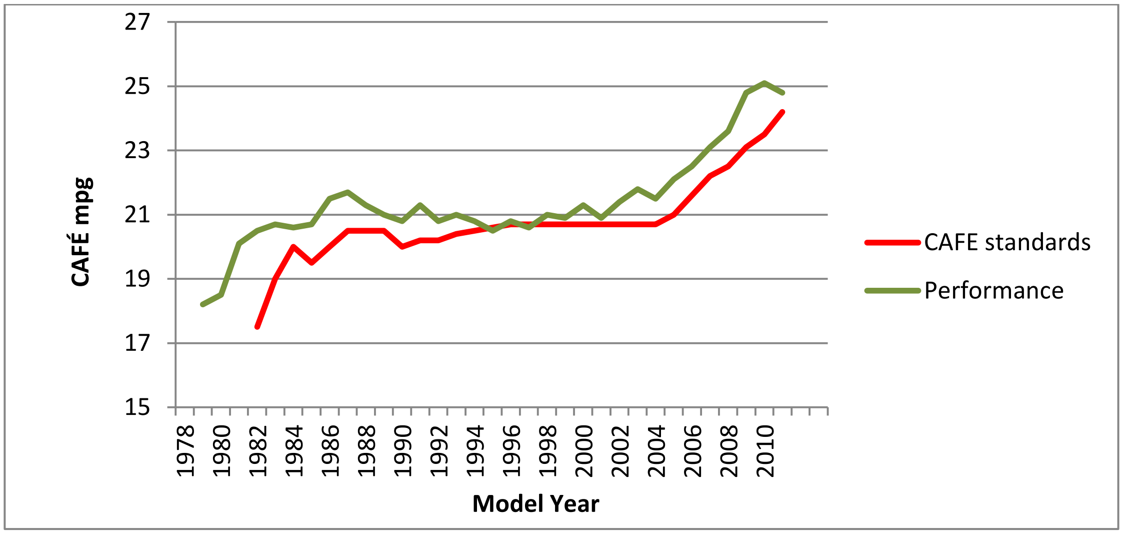

Figure 2-6: Light Truck Fleet Performance, 1978-2011
Data Source: (U.S. Department of Transportation, NHTSA, 2011)

2.3. VMT Fee alternative

2.3.1. Overview

The vehicle-miles-traveled fee option is an alternative that has been implemented in the past in various cases, but not as extensively as it is suggested nowadays. Past implementations of this alternative – mainly in Europe – include a mileage-based fee which varies according to mileage, road and vehicle characteristics, and traffic conditions.

Such a scheme has been implemented in the past but not extensively and homogeneously. Trucks have been included in a VMT scheme in some U.S. states in the past. In Oregon, where the VMT fee revenues of a weight-distance tax reached $178 million in 2003, the scheme includes rates ranging from $0.04 to $0.185 per mile, based on the truck's classification (weight, and number of axles). Under this mileage-based scheme, truck operators report their in- and out-of-state mileage and are exempted from the state fuel tax (TRB Committee for the Study of the Long-Term Viability of Fuel Taxes for Transportation

Finance, 2006). Other studies on the Oregon case include Whitty and Imholt (Whitty & Imholt, 2005), Whitty et al. (Whitty, Svadlenak, & Capps, 2006), and Zhang et al. (Zhang, McMullen, Valluri, & Nakahara, 2009). Other states that have expressed their interest in identifying alternative revenue sources include Alabama (Sisiopiku, Waid, Rizk, McLeod, & Robins, 2006), (Sisiopiku & Waid, 2007), Virginia (Boos & Moruza, 2008), Michigan (Southeast Michigan Council of Governments, 2006) and Iowa (Forkenbrock & Kuhl, 2002).

In Europe as well, there are some countries that have implemented some version of the VMT fee. The German autobahn motorway applies mileage fees to trucks whose weight exceeds the threshold of 12 tons. The toll collection mechanism, with expected revenue of $3 billion per year, has been in effect since January 2005, with an average rate of $0.26 per mile, which varies with number of axles and pollution rating. According to the scheme, 51% of the revenues are to be allocated to road infrastructure, and 49% to railroads and inland waterways. Austria has also implemented mileage fees since 2004, where trucks and buses over 12 tons are subject to tolling. The rate varies from $0.28 to $0.58 per mile, depending on the vehicle type, and all revenues are allocated to motorway infrastructure. In Switzerland, a similar mechanism, which charges vehicles of over 3.5 tons, is in effect since 2001. The rates are higher than those reported for Austria or Germany, as the rate includes an additional charge for externalities such as the environmental effects. Finally, UK plans for a scheme that will charge all drivers according to their VMT, and the actual time and place that their traveling occurs. (TRB Committee for the Study of the Long-Term Viability of Fuel Taxes for Transportation Finance, 2006)

Past research and actual experience have identified some problematic aspects of this mechanism. The two most important issues that have been raised include gaining the public's acceptance, in terms of privacy intrusion, and making the transition from the current mechanism to the new one (TRB Committee for the Study of the Long-Term Viability of Fuel Taxes for Transportation Finance, 2006). The higher the complexity of the implemented VMT fee scheme, the harder the transition shall be, leading to probably higher administrative and enforcement costs. Automation in fee collection shall facilitate the implementation of the mechanism, and may receive greater public acceptance as well. Also, the level of the fee is of great significance as well, as high VMT fees may lead to ''degradation of the system performance and harm the public fare'' (Rufolo, 2011)

2.3.2. Infrastructure and Vehicle Requirements and Cost

Identifying the exact infrastructure and vehicle requirements is not the objective of this thesis. However, providing some general information on the VMT fee mechanism is essential in order to adequately discuss the research question. Past experience is useful in identifying the implementation steps and an estimate of the associated costs for the essential infrastructure and vehicle equipment. Rufolo (2011) provides a comprehensive such overview for various cases where a mileage-based fee is implemented. The main points of his overview are presented in this section.

The FHWA Value Pricing Program suggests various modifications of the mechanism, in all of which a Global Positioning System (GPS) location device is essential in determining the miles traveled (Tollng and Pricing Program: Value Pricing Pilot Program). Such a mechanism is necessary as self-reported mileage is hard to verify and administer. In Oregon, vehicles are equipped with on-board units and the mileage fees are paid at-the-pump, when purchasing fuel. Such a scheme allows for lower administrative and enforcement costs. Information from the on-board units (OBUs) is transmitted to a central station and then to the participating gas stations. The fee collection is realized at the participating gas stations, reducing significantly the administrative and enforcement cost. According to estimations provided by Rufolo (2011), the cost for the necessary hardware and upgrade point-of-sale (POS) systems would be $28.6 million, with an additional $2.7 million for the software and $2.4 million for operating costs at an annual basis. The greater portion of this investment is the OBUs for 3 million vehicles, if implemented in a retrofitting basis; this cost is estimated at approximately $1 billion (Rufolo, 2011).

Peters and Gordon (2009) provide some alternative figures regarding the cost of collection and the OBU cost per vehicle. These numbers include the cost of fee collection at 17.8% of revenue, the OBU cost per vehicle at $125, and an annual operating cost of $0.00179 per VMT (Peters & Gordon, 2009).

In Germany, the tolls are collected via an OBU which allows for reporting toll information to a billing system. In the event that a truck is not equipped with the necessary OBU, payment can be processed online or at a designated toll payment terminal. The initial investment cost by the operator company is estimated at 700 million Euros, with an annual operating cost at 25% of the revenue, which reportedly has been stable over the operation period. The mechanism includes 300 toll checker gantries and 300 enforcement vehicles.

2.4. Past research

Past research has studied various pricing options with regards to their characteristics, advantages and disadvantages, and implications. Due to large number of possible modifications that a pricing option may undergo, the results of past research are not comparable one-to-one across distinct studies. However, the purpose of this thesis is to assess the sustainability performance of the two major rivals in raising revenues for the transportation infrastructure, i.e. fuel tax increases, and the mileage-based fee, in terms of social equity.

2.4.1. Fuel tax per gallon vs. VMT fee: benefits and constraints

The insufficiency of the current sources of revenue is well-known. The question that concerns policymakers and drivers is what kind of reform is more adequate and equitable in order to raise revenues and develop an economically self-sufficient transportation network. The fuel tax per gallon as a source of revenue is not a new suggestion; the benefits and constraints are already known from first-hand experience. Researchers all around the nation have emphasized these advantages and disadvantages in an effort to make the comparison with other pricing alternatives more fathomable.

Pozdena (1995) discusses the financial insolvency of the fuel taxation alternative for the state of California. He comments on the inappropriate use of the existing funds, as a great portion of the revenues collected via the fuel tax pricing option is allocated to mass transit subsidization, instead of undertaking major highway work, such as reconstruction and backlogs on major California roads. He suggests that a congestion pricing scheme may improve the revenue resources and alleviate problematic high traffic volumes (Pozdena, 1995).

Shank & Rudnick-Thorpe (2011) discuss the characteristics of the system, via a pro and con approach. The advantages of maintaining the current funding system include the non-necessity of installing new revenue-collection mechanisms, the non-necessity of extra administrative cost of operating those mechanisms, and the predictability of collected revenues based on historical data. On the other hand, technological advances regarding increased vehicle fuel efficiency, increased and incentive-based use of alternative fuels, and governmental regulations, such as the CAFE standards, reduce the efficiency of the gas tax, making it ineffective in meeting its targets. Additionally, the fuel tax does not correctly represent all the costs associated with the use of the infrastructure; it solely captures the aspect of fuel consumption without accounting for costs such as congestion, or safety, while it fails to successfully capture time- and location-specific characteristics of vehicle use. Still and all, the fact that inflation and increasing construction costs have not been accounted for by policy- and lawmakers, has decreased the purchasing power of the current rate of fuel tax (Schank & Rudnick-Thorpe, 2011).

Litman et al. (1998) conducted a comprehensive multi-page study for the transportation system and the potential revenue sources in the state of Washington. Among various plans, they discuss the pricing option of increasing motor vehicle fuel tax and a mileage-charges plan. Regarding the fuel tax pricing option, Litman et al. (1998) suggest that the maximum increase in fuel tax that would still prevent people from cross-border fuel purchase is $0.20 per gallon. Although this number may be different today, this argument should draw the policymakers' attention to issues such as cross-board fuel purchase. It demonstrates the need for a generic, multi-state policy change, or for regulations that will focus on forestalling such implications. Regarding the plan of mileage-related charges, they view it as a more equitable alternative, in the sense that externalities such as congestion, crashes, emissions, and other, are more correlated with mileage rather than fuel consumption, and also list it as a progressive pricing alternative, with lower income groups being those who drive less, thus will be charged less. Also, they suggest a non-flat per mile charge, in an effort to make the distinction between necessity and luxury. Such a structure would allow a certain number of miles to be taxed in a milder way, as they are considered a necessity; a higher mileage fee is applied towards the next mileage level, being considered as luxury. According to their analysis, the current pricing policies are *‘’unfair, encourage wasteful driving, and limit transportation policies’’*, thus a distance-based system would demonstrate increased horizontal equity across the affected entities. In an effort to identify the best solution for each transportation-related issue, the authors rank the various pricing options. In terms of horizontal equity, mileage-based charges and fuel tax increases rank 2nd and 3rd respectively. With regards to emissions, fuel charges are found to reduce tailpipe emissions, whereas distance-based charges have a larger effect on reducing particulates (Litman, Komanoff, & Howell, Esq., 1998).

Table 2-3: Ranking of fuel tax increases and distance-based charges (Adapted from (Litman, Komanoff, & Howell, Esq., 1998))

Areas of Improvement	**Increase in Fuel Taxes**	**Distance-based charges**
VMT reduction	2nd	1st
Congestion reduction	3rd	2nd
Economic efficiency & horizontal equity	3rd	2nd
Energy conservation	1st	2nd

Areas of Improvement	Increase in Fuel Taxes	Distance-based charges
Vertical equity, with respect to mobility need and ability	3rd	1st

One of the main questions that past research has tried to answer, by drawing information from different datasets and applying different model specifications, refers to the regressivity of fuel taxes across income groups, and its distributional effects with regard to location, demographics, household (HH) life cycle, etc. Poterba (1990) conducted one of the very first studies which studied the regressivity of gasoline tax across income groups. Research methodologies are split between two practices; annual expenditures versus annual income. Poterba (1990) argues that annual expenditures are a more representative measure of the HH's prosperity. By computing elasticities based on annual expenditures rather than annual income, Poterba found fuel tax to be less regressive than previous research did, since the ratio of gasoline-expenditures to annual expenditures is more stable across the population groups than the ratio of gasoline-expenditures to annual income. Using data on annual income and annual expenditures from the 1985 Consumer Expenditure Survey, Poterba (1990) studied both cases in his analysis. In the case where annual income was the grouping variable, households of lower income groups were found to spend more, as much as twice or thrice of their income than HHs of higher income do. In the case of annual expenditures, no particular distributional pattern was found to be in effect. Also, he concludes that, contrary to past research that found fuel tax to be regressive across income groups, the expenditure-based analysis suggests that middle-class households are those who bear the heaviest burden, in terms of gasoline-related expenditures to annual expenditures ratio. More precisely, Poterba's findings provide evidence that the HHs that will be affected the most are those who spend more than 10% of their annual expenditures on gasoline; these HHs are typically located in rural areas and in the South U.S (Poterba, 1990).

Researchers have not limited their study on fuel tax regressivity, but have tried to identify the distributional, equity-wise effects of other pricing options, as well. The state of Oregon is the first state to study the potential of a VMT fee via a pilot study which was initiated in 2006. Although there has been no increase in the state fuel tax of $0.24 per gallon since 1993, still the residents are not willing to support an increase in the aforementioned tax. McMullen et al. (2010) applied both a static and a regression-based model to address the widely-debated issue of substituting the $0.24 per gallon gas tax for light vehicles with a revenue-neutral flat vehicle-mile tax of $0.012 per mile, for the state of Oregon. Their analysis is equity-focused, in terms of HH's location and average income. The static model used, also recommended by the US Congress Joint Committee on Taxation (JCT), concludes that this particular policy change may have a regressive effect, however its size is significantly lower, compared to the effect of fuel price increase over time. Following their analysis, they find that a vehicle-mile tax is slightly more regressive than the fuel tax. Contrary to common expectations HHs located in rural areas benefit more than their counterparts in urban areas (McMullen, Zhang, & Nakahara, 2010).

Shank & Rudnick-Thorpe (2011) discuss the advantages and disadvantages of the VMT fee, which as a pricing option may apply to both federal and state fuel taxes. It is believed to more accurately capture the driver's use of the infrastructure, and users get a better understanding of how their driving behavior affects infrastructure, thus what is the estimated cost of their driving performance. This also allows users to adjust their driving pattern in order to decrease VMT, thus the associated cost. Today, under the prevailing fuel tax, such a linkage between tax paid and travel behavior cannot be intuitively made. Also, the authors consider the VMT fee to be a sustainable alternative, highly transparent, since miles traveled for each vehicle will be tracked and users will demand that their ‘’subsidy’’ contributes visibly to the infrastructure system. However, policymakers are concerned with public acceptance of this potentially new tax, as mileage-tracking devices may be viewed as privacy invaders and also, users who presumably drive more, such as in rural areas, or users with daily commuting, will be harder to persuade. From a practical perspective, implementing a VMT fee will require new revenue-collecting mechanisms that will be likely harder to manage, and would require a transition period of 10-15 years, as discussed in Section 2.3.2 (Schank & Rudnick-Thorpe, 2011).

The University of Iowa conducted a two-year study to assess the technical feasibility and public acceptability of implementing a VMT fee. Hanley & Kuhl (2011) discuss the results of the study that consisted of 2,650 participants in 12 locations across the nation. The technological configuration required installation of an onboard unit (OBU) in each vehicle for 10 months which uses GPS technology to determine location-specific characteristics and assess localized charges. The amount of miles traveled was assessed via either the odometer or the speedometer data of the vehicle. The participants were asked to evaluate the mileage charge; the results suggest that over time the participants became more positive towards the system (moving from 41% favorable to 70% favorable), and only 17% were negative-opinioned at the end of the study. Also, the substantial finding of this study is with regards to privacy; between the maximum privacy configuration and the user auditable configuration, users were more positive-opinioned towards an intermediate level of privacy configuration, namely “modified auditable configuration” (Hanley & Kuhl, 2011).

Weatherford (2011) studies the distributional implications of substituting the federal fuel tax per gallon with a VMT fee of $0.0098 per mile, drawing data from the 2001 National Household Travel Survey. His major findings include:

- VMT fees are less regressive than fuel taxes across income groups, higher income groups being the ones who will bear a heavier tax burden.
- Urban HHs’ burden will be heavier than their rural counterparts’.
- Tax burden will shift from retired HHs to younger HHs with children.

Weatherford’s results suggest that equity should not constitute a concern; on the other hand, privacy issues and high implementation and administrative costs pose major concern. Weatherford emphasizes the advantages of a VMT fee; users are charged the same rate per mile driven on the same road segment. Also, revenue is generated where users actually drive, and not where fuel is purchased, as is the current case. This location-specific attribute of the VMT fee allows for future enhancements/modifications with regards to congestion pricing, a pricing option that may be location- and time-specific, in order to raise revenues respectively to the service provided. On the other hand, he stresses out the inconveniences that are attached to this option; privacy issues, higher possibility for tax evasion, equity concerns and unidentified distributional impacts across social groups. Weatherford discusses some common characteristics of the lower-income groups; these include ownership and operation of older and less fuel efficient vehicles, and higher ratios (40%) of vehicle expenditures to annual expenditures than their high-income counterparts. The equity concerns regarding rural versus urban HHs are based upon data that shows that rural HHs have 16% higher VMT than the average population, and 42% higher VMT than HHs in the urban area. Weatherford's analysis consists of a weighted least squares regression model, based on McMullen et al. (2010). The results indicate that HHs in rural areas, HHs with an average income of $40,000 per year, and HHs that are classified as "retired" with regards to their life cycle phase are those that benefit the most from the VMT fee option. On the other hand, HHs with higher average annual income, HHs in urban and suburban areas, and HHs with children bear a heavier burden. Indicatively, a VMT fee would reduce the annual tax burden of a typical rural HH by $0.57 versus an increase in the annual tax burden of a typical urban HH of $0.79. The main findings of this study show that 87% of the population would experience an increase in the cost-per mile driven of no more than 5%, or 98% of the population would experience an increase of the annual tax burden of no more than $20. Also, with regards to tax revenues location, northeastern and pacific regions would contribute more (Weatherford, 2011).

Additional studies include the Zhang and Lu (2012). They study the marginal cost of a vehicle mileage fee, by taking into account in the model specification externalities such as congestion, and pollution emissions. Their study shows that the implementation of the VMT fee would lead to a significant decrease of 27.1% in VMT, and a slight increase of 4.2% in vehicle fuel efficiency. Energy consumption and pollution emissions were found to decrease by 25% in such case (Zhang & Lu, 2012). Li et al. (2011) studied the effect of gas price changes on vehicle purchases and usage. They found that drivers respond to increasing gas prices by shifting from vehicles of lower fuel efficiency to vehicles of higher fuel efficiency, and by operating their vehicles at lower average speed (Li, Linn, & Muehlegger, 2011).

2.4.2. Other pricing options

2.4.2.1. Feebate and Rebate Programs

Moghadam (2011) discusses whether, and to what extent, incentive-based policies affect vehicle ownership and vehicle operation, and particularly how CO_2 emission levels are affected. Feebate programs mainly affect vehicle ownership, whereas an increase in fuel tax will primarily influence vehicle operation, i.e. miles driven and CO_2 emitted. Also, with regards to reduction of CO_2 levels, the author emphasizes that incentive-based policies do not produce the desired outcome, at a reasonable cost. The author's analysis includes two models; a discrete model on vehicle ownership, and a continuous one vehicle use, i.e. miles driven. His results show that both the likelihood of vehicle ownership and miles driven decrease with increasing gasoline prices, and decreasing average HH income, also, the level of both vehicle ownership and operation is lower in urban areas. Also, higher gasoline price encourages the purchase and use of vehicles of higher fuel efficiency and also decreases the annual demand for miles of 0.036 to 0.26%, depending on the alternative considered. The key element of the structure of any ordinary feebate program is the pivot point, which is the average fuel economy on new vehicles; vehicles with fuel economies that exceed the pivot point receive rebates, whereas vehicles which perform worse than the average fuel-efficiency-wise will be charged fees (Moghadam, 2011).

Similarly to Moghadam (2011), Greene et al. (2005) also discussed the attributes of fuel-economy-driven policies, such as the feebate and rebate programs, and the gas-guzzler taxes. Regarding fuel savings when operating their vehicle, consumers take into consideration only the first three years, however such an arbitrary practice consistently disregards the total fuel savings due to high fuel efficiency during the typical 14-year life cycle of an ordinary vehicle. In a bid to address this misconception and encourage manufacturers and consumers towards the design, purchase and operation of high-fuel efficient vehicles, Greene et al. (2005) argue that the feebate/rebate programs constitute a valid suggestion. They found that a feebate rate of $500 per GPM (gallon per mile) may lead to an increase in fuel economy of as much as 16%. A higher feebate rate would lead to a higher increase in fuel economy (e.g. $1,000 feebate rate – 29% increase). In order to avoid the event where vehicle owners view the feebate programs as another form of taxation, Greene et al. suggest a feebate system that is revenue neutral. Also, to avoid the manufacturers' resentment, they suggest a scheme which has variant feebate rates, according the vehicle's class (Greene, Patterson, Singh, & Li, 2005).

In their study, Greene et al. (2005) also discuss the alternative of gas-guzzler programs. The U.S. guzzler tax program was first enforced in 1989, was then revised in 1991, and it applies to passenger cars with fuel economy of less than 22.5 miles per gallon. It is strictly a fee plan, whereas feebate programs switch from feebate to rebate around the pivot point. Their study emphasizes the risk that a gas-guzzler tax program entails the risk that vehicles will concentrate slightly above the minimum fuel efficiency, in order to avoid paying the tax while keeping the price of the vehicle as low as possible (Greene, Patterson, Singh, & Li, 2005).

2.4.2.2. Energy-related & pollution fees

Krupnick et al. (2001) studied the extent of public support to policies which mainly consist of pollution fees. Their study analyzed the results of a phone survey that was sponsored by REACH Task Force and was conducted in 1996 in Southern California. The underexamination alternative consists of a plan which charges vehicle owners a fee proportional to the vehicle's emission rate (grams per mile*miles driven), but it also includes an alternative module where revenues are recycled through reductions in the sales taxes, vehicle registration fees, or license fees. The latter modification ensured public support from the majority of the interviewees, still the whole plan may be unfair to high mileage drivers (particularly those whose travel includes commuting), and low-income groups, who usually own high-pollutant vehicles to a higher percentage. However, this study underlines the need to be explicit when communicating a new pricing scheme to the public; tangible information on the environmental benefits of the policy, on the destination of the fees collected, as well as ensuring that the fee policy will be as equitable as possible in terms of mileage traveled and vehicle characteristics, will ensure higher levels of public support than past surveys have achieved (Krupnick, Harrington, & Alberini, 2001).

Energy-related fees have been also discussed by Greene (2011). In his study, Greene (2011) discusses the pricing alternative of an indexed energy user fee that would encourage the purchase and operation of more fuel efficient vehicles. According to his findings, "an indexed energy user fee would induce 2 to 4 times as much reduction in GHG emissions and petroleum use as a pure mileage fee", still he identifies its incompetence to manage and potentially reduce congestion and other related issues. Greene (2011) views the indexed energy user fee (Indexed Roadway User Toll on Energy – IRoUTE) as the next reasonable pricing option in financing the US surface transportation, while ensuring that we move towards more sustainable practices, particularly environmentally focused. His suggestion is highly driven by the fact that *"highway vehicles generate 25% of US* CO_2 *emissions and account for over 70% of its petroleum consumption."* Comparing an indexed road user toll on energy to the oft-proposed VMT fee, Greene (2011) concludes that it encourages environmentally friendly practices, since in the case of an increase in energy price, users are driven towards purchasing more fuel efficient vehicles, as well as operate them in a more conservative and environmentally friendly way. The issue of loss of purchase power over time can be addressed successfully by indexing the tax to inflation. Additionally the author argues that such an alternative would have a lower administrative cost than switching to a VMT fee, an advantage which renders it quite competitive and ready for implementation. However he points out the fee's inability to manage congestion, promote the use of alternative fuels, as well as weigh appropriately the impact of heavy vehicles on the infrastructure (Greene, 2011).

West (2004), in her analysis, focused more on alternative vehicle pollution control policies, such as subsidization policies that favor new vehicles versus older, and usually less fuel-efficient ones. Past research has shown that such policies successfully lower vehicle-related pollution. In general, gas-guzzler taxes, CAFE standards, accelerated vehicle retirement programs encourage the purchase and operation of fuel efficient, small and new cars respectively. However, regarding the progressivity/regressivity of each alternative across income groups, past research has shown that results may vary, depending on the particular measure of "income" and vehicle ownership. Some studies consider annual income and exclude zero-vehicles HHs, whereas other studies consider consumption expenditures and include in their analysis non-vehicle HHs as well. In the former case, all policies are found to be regressive across all income groups, but in the latter one, gas tax is found to be less regressive. West (2004) estimates price elasticities and simulates changes in mileage in a two-step analysis. The first step consists of a nested logit model for the HH's vehicle choice with regards to number of vehicles, vintage and engine size. The second step models the demand for VMT. The results of West's analysis emphasize that both the VMT fee and the fuel tax are regressive only across upper income groups; this is justified by the fact that lower income groups are more likely to belong to the non-vehicle category, and also they tend to reduce their VMT more extensively, in the event of a price increase (West, 2004).

Lindsey et al. (2011) studied the effect of residential location on VMT, energy consumption, and GHG emissions, using data for 2007-2008 from the Chicago metropolitan area. They found that increasing residential distance from the city center and decreasing residential density lead increased VMT. They also found that shifting from lower to higher vehicle fuel efficiency will be more effective in reducing GHG emissions than reducing VMT. The last finding is also supported by the scenario that adopting the 2012 European fuel economy standards causes a 48% decrease in GHG emissions (Lindsey, Schofer, Durango-Cohen, & Gray, 2011).

Ensuring public support is the key element that will make a policy successful. Various surveys have been conducted in an effort to identify if the public supports or is opposed to various pricing options. Deakin & Harvey (1996) developed a very extensive report that discusses five categories of transportation pricing measures, namely congestion pricing, parking charges, fuel tax increases, VMT fees, and emission fees. Apart from the actual discussion on the five different pricing options, they also commented on public acceptability of these options. They suggest that the implemented strategy should match the regional conditions, in order to achieve public acceptance. They found that people are somehow willing to pay more if the increased revenues are devoted to the transportation system, and are not allocated to other directions. They found that people support a fuel tax increase rather than any of the other four alternatives. Particularly for the gas tax increase option, people view it as a mean to decrease VMT in the short-run and shift from lower to higher vehicle fuel efficiency in the long run. The VMT fee was found to be a good suggestion in the event of alternative vehicles (such as electric cars) achieve a higher market penetration (Deakin & Harvey, 1996).

In an effort to identify the extent of public's support to practices which aim to reduce travel-related air-pollution, and U.S. dependency on oil imports, and which factors affect this extent, Agrawal et al. (2010) conducted and analyzed a phone-survey among California residents in a bid to identify whether they would support a policy scheme that is environmentally-driven, via the implementation of "green" transportation taxes and fees. In 2008, 1,200 Californians were asked to evaluate five policy alternatives which aim at maintaining and improving the transportation network, namely a flat registration fee increase of $31 per vehicle, a green vehicle registration fee increase (varying relatively to the pollution induced), a feebate, a flat mileage fee of $0.01 per mile within the state), and a green mileage fee (varying relatively to the pollution induced). Their two-step analysis, which consisted of a bivariate and a multivariate analysis, yielded some very interesting results The highest acceptance rates were received by the feebate proposal and the potential replacement of the gas tax by a variable mileage fee. Also, 64% of the interviewees are positive towards a green vehicle registration fee, under the premise that the revenues will be allocated to environmentally-driven transportation programs. As to the factors that may affect public reaction to policy alternatives that focus on reducing travel-induced pollution. Sex, age, and level of education have an impact on willingness-to-pay (women, young people, and more educated people demonstrate a more positive attitude towards such policies). Agrawal et al. (2010) conclude that, environmental knowledge is a good predictor of the support towards green products, thus policymakers may need to focus on educating the public of the environmental costs and benefits of various policy alternatives (Agrawal, Dill, & Nixon, 2010).

2.4.2.3. Dedicated General Revenue

The option of dedicated general revenue is discussed by Shank & Rudnick-Thorpe (2011). Similarly to any other domestic discretionary program, transportation funding will be a standard percentage of the general revenues, allowing consistent funding over time. It also assumes dissolution of the HTF. This revenue option does not require complex revenue-collecting mechanisms, and the long-term competition between highway and transit and among various states to ensure higher shares of federal funding will be mitigated. Still, emphasis should be put on the fact that transportation projects take years to complete, so funding should be dedicated to allow for such projects to be undertaken (Schank & Rudnick-Thorpe, 2011).

2.5. Summary

The preceding discussion shows that there has been some research in the broader context of pricing options, in an effort to identify alternative revenue sources, but also in a bid to mitigate some of the negative effects of high traffic volumes. The current fuel tax pricing option has been debated to be insufficient in terms of necessary revenues to maintain the surface transportation system. A vehicle-miles traveled fee pricing option has been examined in terms of infrastructure and vehicle requirements for implementation, the associated costs, but also in terms of its regressivity among groups.

Past research has motivated the author to explore the research question further and identify which social groups are most likely to be affected under the two pricing options studied. The equity performance of each alternative is studied at the national level, contrary to previous studies which have focused to a more regional level, in terms of experimental design and subsequent statistical analysis.

Chapter 3: Data Description discusses the data used in the statistical analysis, in terms of data processing, and data description.

Chapter 3: Data Description

3.1. Introduction

The dataset used for the current analysis is Version 2.1 of the 2009 National Household Travel Survey (2009 NHTS), which was updated in February 2011 (U.S. Department of Transportation, Federal Highway Administration). Conducted between March 2008 and March 2009, the 2009 National Household Travel Survey *"serves as the nation's inventory of daily travel"* and the information provided is aimed *"to assist transportation planners and policy makers who need comprehensive data on travel and transportation patterns in the United States"* (U.S. Department of Transportation, Federal Highway Administration).

The 2009 NHTS includes useful information on demographic-, socioeconomic-, and travel-related variables. It consists of four distinct dataset files; the Household File, the Person File, the Vehicle File, and the Day-trip File. Herein, the level of analysis is the household (HH); however, the need to include both HH- and vehicle-specific characteristics in this analysis suggested the need to merge the two distinct datasets into one consistent file. The final dataset is at the HH level, but it also provides aggregated information on vehicle characteristics (U.S. Department of Transportation, 2009 National Household Travel Survey: User's Guide (Version 2)), (U.S. Department of Transportation, 2009 National Household Survey: Codebook).

Since the research objectives of this thesis include addressing the same research question at two distinct levels of analysis (national vs. local), the discussion on the data characteristics follows this pattern. This chapter follows the following structure; Section 3.2 describes the procedure that the author followed in order to successfully merge the two distinct data files into one final and consistent file at the HH level. Section 3.3 provides information on the procedure of applying the essential weights, in order to make the sample representative of the population. Section 3.4 describes the process the author followed in order to remove some observations which were considered as ‘’outliers’’. Section 3.5 provides information on the descriptive statistics of the finalized dataset, both at the national and the local (Iowa) level. Section 3.6 provides comparative plots of the variables of interest between the national and the local model, presented under the pattern of the two pricing options compared in this thesis, in order to facilitate the presentation of results and discussion in Chapter 4: Methodology and Results. Finally, Section 3.7 summarizes the key conclusions yielded from this chapter.

3.2. Merging the Household and the Vehicle Files

The nature of the research question required the use of both HH- and vehicle-specific characteristics in the model development. The two files (Household File and Vehicle File) have some inherent differences that should be accounted for when merging them into one file.

The original Household File consists of 150,147 observations (one observation per HH) and it includes variables which could be broadly categorized as follows:

- ***Location-specific variables***, such as Census region classification for HH home address, State Federal Information Processing Standard (FIPS) for HH home address, MSA population size for the HH home address.
- ***Demographics***, such as life cycle classification for the HH, race of HH respondent, count of adult HH members at least 18 years old.

- ***Socioeconomic variables***, such as housing unit owned or rented, derived total HH income, type of housing unit.

The original Vehicle File consists of 309,163 observations (one observation per vehicle) and it includes variables which are vehicle-specific such as the EIA derived miles per gasoline-equivalent gallon estimate, the vehicle type, how long the vehicle is owned by the HH (in months), vehicle age, or the best estimate of annual miles. In order to be consistent with past research, the level of the analysis conducted in this thesis is the HH level (Poterba, 1990), (Zhang, McMullen, Valluri, & Nakahara, 2009), (Moghadam, 2011), (McMullen, Zhang, & Nakahara, 2010), (Litman, 1999), (Lindsey, Schofer, Durango-Cohen, & Gray, 2011), (Kayser, 2000), (Fullerton & West, 2002), (Busse, Knittel, & Zettelmeyer, 2009), (Bento, Goulder, Jacobsen, & von Haefen, 2009), (West, 2004), (Larsen, Burris, Pearson, & Ellis, 2012). Also, the author has excluded 7,205 observations of HHs with zero vehicles from the Household File, as these HHs did not generate observations on fuel efficiency and VMT that are of interest in this study (McMullen, Zhang, & Nakahara, 2010).

The merging process was performed in Microsoft Access 2010 on the basis of the HH's unique eight-digit ID number (HOUSEID variable). The HHVEHCNT variable contains information regarding how many vehicles the HH owns, thus how many observations from the Vehicle File were merged into one for the final dataset at the HH level.

3.2.1. National data

Following the merging process, the final table for the national model consists of 142,942 observations. As this analysis is conducted at the HH level, there was a need to generate a single vehicle observation per HH, even if the HH may own more than one vehicle. For that case, the author either used the sum (S) or the weighted average (WA), depending on the nature of the variable; this piece of information is indicated in the last column of Table 3-2. The weighted average was computed based on the best estimate of annual miles (BESTMILE). The following equation serves as an example of how weighting according to BESTMILE was applied in order to compute the weighted average of the vehicle fuel efficiency at the HH level:

$$WEIADMPG = \frac{\sum_{i=1}^{n}(EIADMPG_i \times \ BESTMILE_i)}{\sum_{i=1}^{n} BESTMILE_i}$$

According to this process, the average vehicle fuel efficiency at the HH level was computed based on the mileage of each vehicle owned by the HH, to account for vehicle usage.

3.2.2. Local data

Identifying the socioeconomic and demographic groups which will be mostly impacted under each pricing option is addressed both at the national and local level. At the nationwide level, the author uses the dataset of 142,942 observations, however it is expected that, due to the large amount of information contained in the dataset, the models developed in Chapter 4: Methodology and Results will be able to explain little of a lot of variability (expected R^2 is low). For the local model, the author studies the same research question for the state of Iowa. The reason of this research interest is the objective to compare the results at two different levels of analysis.

Figure 3-1: Iowa on the U.S. map (Vacation 2 USA)

The dataset for the local model (Iowa) was extracted from the national data, based on the HH's FIPS coding (HHSTFIPS variable). For Iowa, the HHSTFIPS variable is set equal to 19. Following the filtering process, the final table for the local model consists of 3,614 observations. Since the local data originated from the national dataset through filtering, the vehicle attributes are either the sum (S) or the weighted average (WA) of the corresponding variable for a unique vehicle observation of the original 2009 NHTS Vehicle table, depending on the nature of the variable. This piece of information is indicated in the last column of Table 3-4.

3.3. Applying the weights

Weighting the data is a critical process that needs to be performed so that the modeling process yields realistic results. In the interest of accounting for non-response, under-coverage, and multiple telephones in a HH, the data is weighted by the final HH weight WTHHFIN provided both in the Household and Vehicle File (Rizzo, et al., 2009). Since vehicle ownership is estimated at a HH basis, both the original HH and Vehicle tables include the same – at the HH level – weighting variable, i.e. WTHHFIN.

3.4. Removing the outliers

Upon merging the HH and the vehicle table, the final number of observations is 142,942. Not all variables, as listed in Table 3-2 and Table 3-4, have reported values for all 142,942 for the national model and 3,614 for the local. In Table 3-2 and Table 3-4, the fifth column indicates the number of reported and the number of missing values for each variable. However, a detailed descriptive statistics table for the variables that were originally provided in the two datasets indicated that some observations may be considered as outliers, and should not be included in our analysis. Even though the sample is fairly large, it is still highly probable that these outliers may drive the model results. Table 3-1 shows the range of the problematic variables before and after removing the outliers. Also, the percentages in parentheses infer that less than 3.87% of the original information is not considered in the development of the models.

Table 3-1: Value range before and after removing the outliers (with associated percentages of the total number of observations)

Original Variable Mnemonic	**New Variable Mnemonic**	**Variable Description**	**Original range (% of the sample)**	**New range - outliers removed (% of the sample)**
DRVRCNT	DRVRCNT2	Number of drivers in HH	0-9 (100%)	0-4 (99.62%)
HHSIZE	HHSIZE2	Count of HH members	1-14 (100%)	1-6 (99.30%)
HHVEHCNT	HHVHCNT2	Count of HH vehicles	1-27 (100%)	1-6 (99.52%)
NUMADLT	NUMADLT2	Count of adult HHMs at least 18 years old	1-10 (100%)	1-4 (99.56%)
WRKCOUNT	WRKCNT2	Number of workers in HH	0-6	0-3

Original Variable Mnemonic	New Variable Mnemonic	Variable Description	Original range (% of the sample)	New range - outliers removed (% of the sample)
			(100%)	(99.39%)
N_VEHCOM	N_VEHCO2	Number of commercial license plate vehicles in HH	0-11 (100%)	0-2 (99.74%)
N_AUTO	N_AUTO2	Number of automobile/car/station wagon vehicles in HH	0-22 (100%)	0-3 (99.22%)
N_SUV	N_SUV2	Number of sport utility vehicles in HH	0-10 (100%)	0-2 (99.54%)
N_PICKUP	N_PICUP2	Number of pick-up truck vehicles in HH	0-9 (100%)	0-2 (99.21%)
N_OTHER	N_OTHER2	Number of other vehicles in HH	0-27 (100%)	0-6 (99.49%)
N_MGAS	N_MGAS2	Number of motor gasoline vehicles in HH	0-26 (100%)	0-6 (99.55%)
VEHOWNMO	VEHOWNM2	How long vehicle(s) owned in months	0-623.211 (100%)	0-186.963 (98.39%)
VEHAGE	VEHAGE	Age of vehicle in years	1-35 (100%)	1-24 (96.13%)
WEIADMPG	CWEIADMP	EIA derived miles per gasoline-equivalent gallon estimate (weighted average)	6.4-117 (100%)	10.88-32.18 (96.28%)

Additionally, in order to identify whether these outliers are related to each other, and whether, for each outlier, the whole observation should be dropped from the analysis sample, the author estimates the correlation matrix among these variables in Appendix 1C: Correlation Matrix for Outliers. The results infer that there is no particular relationship among the outliers of two or more variables. Finally, the validity of the responses has been established by computing the differences between independent variables and determining the sign; for example, across all observations, the reported HH size is greater than the reported number of drivers per HH and the reported number of HHMs at least 18 years old.

Particularly for the vehicle fuel efficiency variable WEIADMPG, the range of the variable (6.4 to 117 miles per gasoline-equivalent gallon estimate) is notably wide. In order to avoid strange data behavior at the limit values of this particularly wide range, the author suggests re-defining the limits of the range, by creating a new variable (CWEIADMP) which consists only of the values which are within two standard deviations of the mean of the original WEIADMPG variable.

$$(21.5297 - 10.6466, 21.5297 + 10.6466) = (10.8831, 32.1763)$$

Upon confining the range of the variable, the information still represents 96.28% of the information provided by the original variable.

3.5. Descriptive Statistics

3.5.1. National data

Table 3-2 provides the weighted descriptive statistics of the variables involved in the final national model specification. A detailed table summarizing the descriptive statistics of all variables considered in the national model development process is presented in Appendix 1A: Descriptive Statistics – Nation. Additionally, Table 3-3 presents the correlation matrix of the variables included in the final specification for the national models. The information of this table has been very useful for checking purposes of the qualitative results of the model development, in terms of signs. The correlation matrix of all the variables considered in the model development is included in Appendix 2A: Correlation Matrix – Nation.

Note that, in the second column of Table 3-2 and Table 3-4, along with the variable description, the following pieces of information are also provided:

- [corrected values]: indicates that the outliers have been removed from the original variable, according to the procedure described in Section 3.4.
- *log:* indicates that the variable is logged.
- [continuous]: indicates that the variable is continuous.
- [dummy]: indicates that it is a dummy variable (values: 0, 1).
- [count]: indicates that it is a count variable.

Table 3-2: Weighted descriptive statistics – National Data

Variable Mnemonic	**Variable Description**	**Mean (Std. Dev.)**	**Min/Max**	**Cases (missing)**	**A, S, or WA**
LEFFC	EIA derived miles per gasoline-equivalent gallon estimate [corrected values] – [continuous] – *log*	3.02616 (0.199201)	3.38799/ 3.47117	122,703 (20,239)	WA
LVMT	Best estimate of annual miles [corrected values] – [continuous] - *log*	9.57729 (1.00737)	3.56137/ 12.9215	138,056 (4,886)	S
LARGE	Vehicle type: Pickup truck or SUV [dummy]	0.51471 (0.499785)	0/1	142,942 (0)	S
WRKTOSIZ	Number of workers in HH to HH size ratio – [corrected values] - [continuous]	0.502896 (0.378755)	0/1	140,141 (2,801)	A
VEHOWNM2	How long vehicle owned – months [corrected values] – [continuous]	51.7591 (36.874)	0/186.95	130,565 (12,377)	WA
VEHAGE	Age of vehicle in years [corrected values] – [continuous]	7.58189 (4.55294)	0/24	137,902 (5,040)	WA
HOMEOWN	Housing unit owned [dummy]	0.71289 (0.452415)	0/1	142,942 (0)	A
RAIL	MSA heavy rail status for HH [dummy]	0.263987 (0.440794)	0/1	142,942 (0)	A
TAX	Federal and state fuel tax per gallon [continuous]	0.495559 (0.10973)	0.264/ 0.68	131,871 (11,071)	A
URB	Household in urban area [dummy]	0.757218 (0.428766)	0/1	131,870 (11,072)	A
NOSUB	HH owns only one vehicle [dummy]	0.353514 (0.478062)	0/1	142,942 (0)	S
ONEADULT	HH life cycle classification: one adult [dummy]	0.286721 (0.452232)	0/1	142,942 (0)	A
LAND	Land use: residential [dummy]	0.873524	0/1	142,942	A

Variable Mnemonic	Variable Description	Mean (Std. Dev.)	Min/Max	Cases (missing)	A, S, or WA
		(0.332386)		(0)	
HYBRID	Vehicle is hybrid or uses alternate fuel [dummy]	4.31E-02 (0.203052)	0/1	142,942 (0)	S
DRVRCNT2	Number of drivers in HH [corrected] – [count]	1.8229 (0.746648)	0/4	142,395 (547)	A
LINCOME	Derived total HH income [continuous] - *log*	10.65 (0.811549)	7.82405/ 11.5129	131,871 (11,071)	A
LMSASIZ	MSA population size for the HH home address [corrected] – [continuous] - *log*	11.247 (5.77538)	0/14.9141	131,870 (11,072)	A

Table 3-3: Correlation matrix – National data

	HOMEOWN	RAIL	VEHAGE	DRVRCNT2	VEHOWNM2	URB
HOMEOWN	1	-0.1548	0.0267	0.0970	0.1097	-0.1177
RAIL	-0.1548	1	-0.0330	0.0265	-0.0028	0.1744
VEHAGE	0.0267	-0.0330	1	0.1412	0.6071	-0.0428
DRVRCNT2	0.09704	0.0265	0.1412	1	0.0374	-0.0349
VEHOWNM2	0.1097	-0.0028	0.6071	0.0374	1	-0.0042
URB	-0.1177	0.1744	-0.0428	-0.0349	-0.0042	1
ONEADULT	-0.1604	0.0254	-0.1175	-0.6454	-0.0750	0.0699
NOSUB	-0.2049	0.0413	-0.2398	-0.5388	-0.1711	0.1057
LAND	0.0338	0.0129	0.0117	0.0284	0.0024	-0.1346
TAX	-0.0336	0.3523	-0.0147	0.0069	0.0007	0.0705
WRKTOSIZ	-0.1179	0.0803	0.0159	0.1045	-0.0391	0.0260
LINCOME	0.1533	0.0919	-0.0544	0.2910	0.0255	0.0414
LMSASIZ	-0.0469	0.2937	-0.0486	0.0221	0.0071	0.3600
LVMT	0.0738	-0.0141	0.0462	0.4517	-0.0225	-0.1278
LEFFC	-0.0884	0.0790	-0.1172	0.0542	-0.1243	0.0640
HYBRID	0.0096	-0.0034	-0.0471	0.0226	-0.0439	-0.0222
LARGE	0.1239	-0.0853	0.0692	0.2444	0.0392	-0.1709

	ONEADULT	NOSUB	LAND	TAX	WRKTOSIZ	LINCOME
HOMEOWN	-0.1604	-0.2049	0.0338	-0.0336	-0.1179	0.1533
RAIL	0.0254	0.0413	0.0130	0.3523	0.0803	0.0919
VEHAGE	-0.1175	-0.2400	0.0117	-0.0148	0.0159	-0.0544
DRVRCNT2	-0.6454	-0.5388	0.0284	0.0069	0.1045	0.2910
VEHOWNM2	-0.0750	-0.1711	0.0024	0.0007	-0.0391	0.0255

	ONEADULT	NOSUB	LAND	TAX	WRKTOSIZ	LINCOME
URB	0.0699	0.1057	-0.1347	0.0705	0.0260	0.0414
ONEADULT	1	0.6067	-0.0312	0.0057	0.0087	-0.3046
NOSUB	0.6067	1	-0.0340	0.0284	-0.1158	-0.3557
LAND	-0.0312	-0.0340	1	0.0186	-0.0069	0.0044
TAX	0.0057	0.02836	0.0186	1	-0.0164	0.0264
WRKTOSIZ	0.0087	-0.1158	-0.0069	-0.0164	1	0.2590
LINCOME	-0.3046	-0.3557	0.0044	0.0264	0.2590	1
LMSASIZ	0.0033	0.0361	-0.0020	0.1200	0.0221	0.1363
LVMT	-0.3702	-0.4906	0.0401	-0.0428	0.2676	0.3547
LEFFC	0.0197	0.0413	-0.0083	0.0678	0.1339	0.0381
HYBRID	-0.0265	-0.0381	0.0067	-0.0148	0.0046	0.0174
LARGE	-0.2541	-0.3980	0.0451	-0.0763	0.1045	0.1715

	LMSASIZ	LVMT	LEFFC	HYBRID	LARGE
HOMEOWN	-0.0469	0.0738	-0.0884	0.0096	0.1239
RAIL	0.2937	-0.0141	0.0790	-0.0034	-0.0853
VEHAGE	-0.0486	0.0462	-0.1172	-0.0471	0.0692
DRVRCNT2	0.0221	0.4517	0.0542	0.0226	0.2444
VEHOWNM2	0.0071	-0.0225	-0.1243	-0.0438	0.0392
URB	0.3600	-0.1278	0.0640	-0.0222	-0.1709
ONEADULT	0.0033	-0.3702	0.0197	-0.0265	-0.2541
NOSUB	0.0361	-0.4906	0.0413	-0.0381	-0.3980
LAND	-0.0020	0.0401	-0.0083	0.0067	0.0451
TAX	0.1200	-0.0428	0.0678	-0.0148	-0.0763
WRKTOSIZ	0.0221	0.2676	0.1339	0.0046	0.1045
LINCOME	0.1363	0.3547	0.0381	0.0174	0.1715
LMSASIZ	1	-0.0509	0.0643	-0.0160	-0.1150
LVMT	-0.0509	1	0.2662	0.0370	0.3038
LEFFC	0.0643	0.2662	1	0.0071	-0.3231
HYBRID	-0.0156	0.0370	0.0071	1	0.0515
LARGE	-0.1150	0.3038	-0.3231	0.0515	1

3.5.2. Local data

Table 3-4 provides the weighted descriptive statistics of the variables involved in the final model specification. A detailed descriptive statistics table of all variables considered in the model development process is presented in Appendix 1B: Descriptive Statistics – Iowa. Also,

Table 3-5 displays the correlation matrix of the variables considered in the final model development at the local level. The correlation matrix of all the variables considered in the model development is included in Appendix 2B: Correlation Matrix – Iowa.

Table 3-4: Weighted descriptive statistics: Local data (Iowa)

Variable Mnemonic	Variable Description	Mean (Std. Dev.)	Min/Max	Cases (Missing)	A, S, or WA
LVMT	Best estimate of annual miles [corrected values] – [continuous] - *log*	9.69461 (1.01012)	0/12.0931	3,491 (123)	S
LEFFC	EIA derived miles per gasoline-equivalent gallon estimate [corrected values] – [continuous] - *log*	3.0021 (0.1836)	2.38876/ 3.47116	3,170 (444)	WA
HYBRID	Vehicle is hybrid or uses alternate fuel [dummy]	6.40E-02 (0.244788)	0/1	3,614 (0)	S
VEHOWNM2	How long vehicle owned – months [corrected values] – [continuous]	51.7811 (36.6158)	0/186.843	3,294 (320)	WA
VEHAGE	Age of vehicle in years [corrected values] – [continuous]	7.93991 (4.16266)	0/24	3,488 (126)	WA
LARGE	Vehicle type: Pickup truck or SUV [dummy]	0.569418 (0.495226)	0/1	3,614 (0)	S
WRKTOSIZ	Number of workers in HH to HH size ratio – [corrected values] - [continuous]	0.540065 (0.385218)	0/1	3,536 (78)	A
LMSASIZ	MSA population size for the HH home address [corrected] – [continuous] - *log*	5.63116 (6.11296)	0/13.5278	3,367 (247)	A
NOSUB	HH owns only one vehicle [dummy]	0.255191 (0.436029)	0/1	3,614 (0)	S
URB	Household in urban area [dummy]	0.579106 (0.493775)	0/1	3,367 (247)	A
LINCOME	Derived total HH income [continuous] - *log*	10.649 (0.722012)	7.82405/ 11.5129	3,367 (247)	A
ONEADULT	HH life cycle classification: one adult [dummy]	0.281483 (0.449785)	0/1	3,614 (0)	A

Table 3-5: Correlation matrix – Local data (Iowa)

	VEHAGE	VEHOWNM2	URB	ONEADULT	NOSUB	WRKTOSIZ
VEHAGE	1	0.5310	-0.0481	-0.1232	-0.2360	0.0320
VEHOWNM2	0.5310	1	0.0233	-0.0425	-0.1322	-0.0459
URB	-0.0481	0.0233	1	0.0470	0.1187	-0.0722
ONEADULT	-0.1232	-0.0425	0.0470	1	0.6202	-0.0326
NOSUB	-0.2360	-0.1322	0.1186	0.6202	1	-0.1264
WRKTOSIZ	0.0320	-0.0459	-0.0722	-0.0326	-0.1264	1
LINCOME	-0.0570	-0.0284	-0.0086	-0.3555	-0.3634	0.2872
LMSASIZ	-0.0628	0.0186	0.3952	-0.0291	0.0465	-0.0747
LVMT	0.0288	-0.1036	-0.2185	-0.3569	-0.4569	0.2644
LEFFC	-0.0945	-0.1064	0.0638	-0.0286	0.0154	0.0640
HYBRID	-0.0579	-0.0210	-0.0502	0.0225	0.0209	0.0404
LARGE	0.0127	-0.0192	-0.2185	-0.2206	-0.3734	0.1577

	LINCOME	LMSASIZ	LVMT	LEFFC	HYBRID	LARGE
VEHAGE	-0.0570	-0.0628	0.0288	-0.0945	-0.0579	0.0127
VEHOWNM2	-0.0284	0.0186	-0.1036	-0.1064	-0.0210	-0.0192
URB	-0.0086	0.3952	-0.2185	0.0639	-0.0502	-0.2185
ONEADULT	-0.3555	-0.0291	-0.3569	-0.0286	0.0225	-0.2206
NOSUB	-0.3634	0.0465	-0.4569	0.0154	0.0209	-0.3734
WRKTOSIZ	0.2872	-0.0747	0.2644	0.0640	0.0404	0.1577
LINCOME	1	0.1107	0.3728	0.0562	0.0312	0.1712
LMSASIZ	0.1107	1	-0.1187	0.0938	-0.0274	-0.1687
LVMT	0.3728	-0.1187	1	0.2575	0.0078	0.3090
LEFFC	0.0562	0.0938	0.2575	1	-0.0993	-0.2840
HYBRID	0.0312	-0.0274	0.0078	-0.0993	1	0.0324
LARGE	0.1712	-0.1687	0.3090	-0.2840	0.0324	1

3.6. Plotting & interpreting the data

Interpreting the data is essential before developing the models which would adequately address the research question. Under each pricing option (fuel tax vs. VMT fee), the author identifies which variables of the dataset may serve as good representatives of each alternative. Then, the author plots them stratified by various levels of the independent variable(s) X.

For the VMT fee, the choice was straight-forward; the dependent variable is the best estimate of annual miles (BESTMILE). In the case of the fuel tax alternative, the choice of the dependent variable is a more challenging process; the three (3) dependent variables which were initially chosen as good representatives of the fuel tax alternative are:

- EIA derived miles per gasoline-equivalent gallon estimate (weighted average),
- Annual fuel expenditures in nominal US dollars (sum), and
- Annual fuel consumption in gasoline equivalent gallons (sum).

The final selection was the EIA derived miles per gasoline-equivalent gallon estimate (weighted average), as it is considered that it best captures the HH sub-groups which are affected at different levels by the current pricing option of the fuel tax per gallon.

3.6.1. Fuel Tax – Y: Average vehicle fuel efficiency at the HH level

The following graphs show the relationship between the dependent variable (EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values], i.e. the average vehicle fuel efficiency at the HH level) with some of the variables of the data. Plotting the variables is helpful in identifying which factors are most likely to affect the dependent variables, and also, together with the correlation matrices, are valuable in determining the expected sign of the independent variables' coefficients in the model development of Chapter 4: Methodology and Results.

Figure 3-2 shows the relationship between the average vehicle fuel efficiency and the home ownership status. Respondents who own the home they reside in tend to have vehicles of lower fuel efficiency, both at the nationwide and the local (Iowa) level. It is possible that people cannot afford to purchase both a housing unit and a vehicle of higher fuel efficiency, but there may be other underlying factors which affect this relationship, and which will be identified in more detail via the model development.

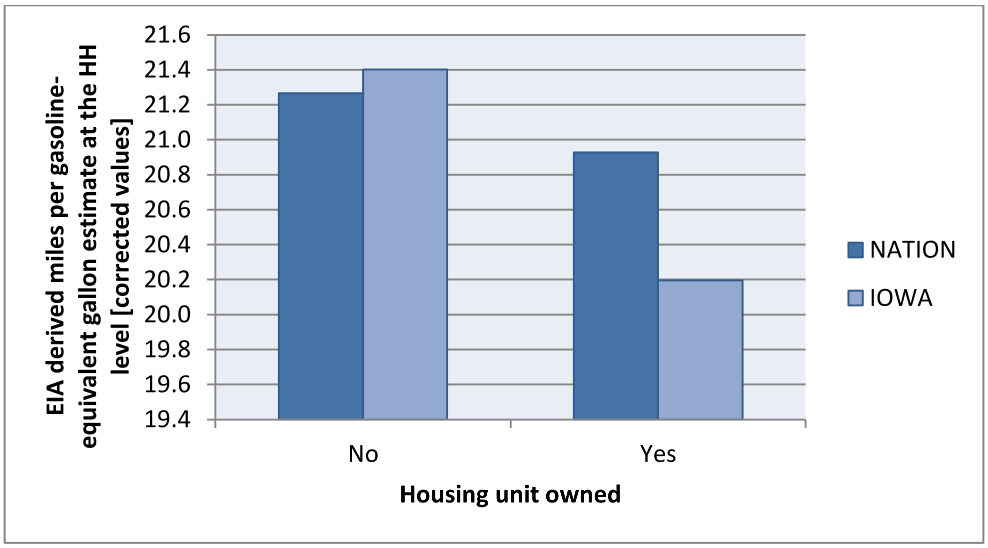

Figure 3-2: EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values] by housing ownership status: Nation vs. Iowa

Accessibility to rail is a factor which affects the average fuel efficiency of the vehicle owned by the HH. Contrary to what may be expected, both at the national and regional level, HHs which have access to rail still own more fuel efficient vehicles, even though rail serves as a mode alternative to personal vehicle (Figure 3-3). This difference, though, is slight, and it should be noted that for the local level (Iowa) the RAIL variable is always equal to 0, thus cannot be included in the model development as it displays no variation.

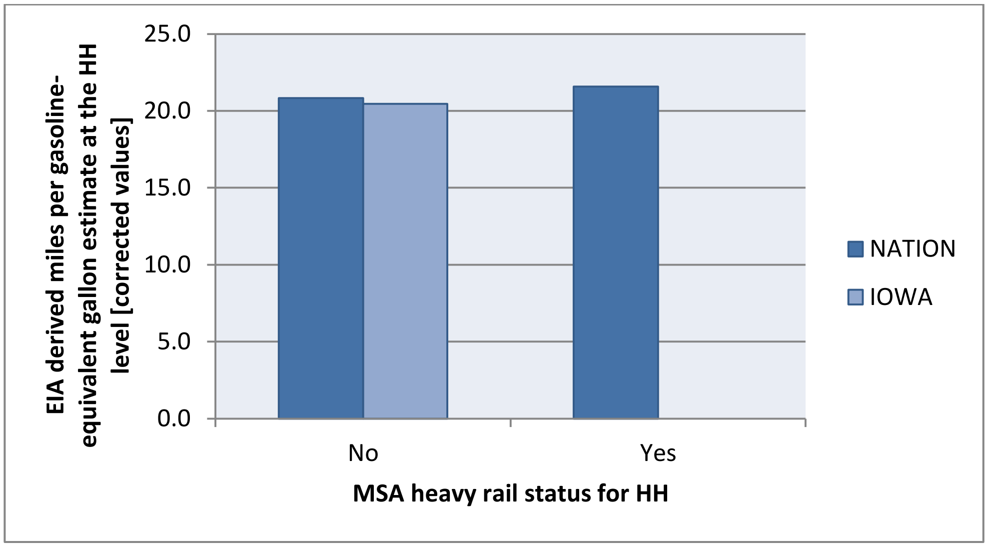

Figure 3-3: EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values] by MSA heavy rail status for HH: Nation vs. Iowa

Figure 3-4 shows how average vehicle fuel efficiency at the HH level varies, with regard to the HH life cycle classification. Although there is no strong pattern identified in this graph, HHs which consist of only one adult seem to display common behavior. For the national dataset, it is found that the subcategory of one adult, with youngest child 0-5 years old, operates vehicle(s) of higher fuel efficiency than any other subgroup. For the local data, this observation is valid for the 2+ adults, youngest child 6-15 subgroup, although other subgroups share values in the close proximity. It is evident though that for both the national and the local datasets, the average vehicle fuel efficiency at the HH level drops for the HH whose member(s) are retired.

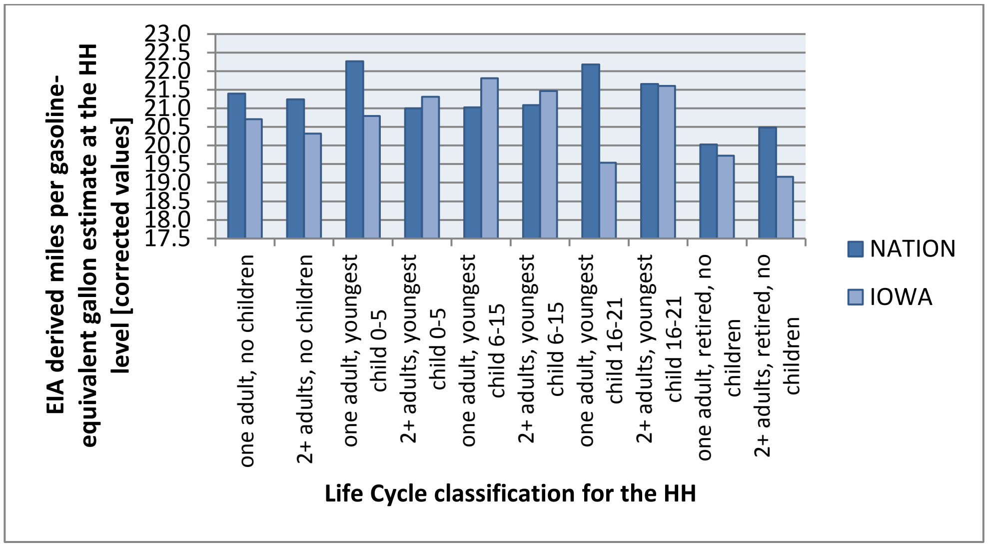

Figure 3-4: EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values] by life cycle classification for the HH: Nation vs. Iowa

The relationship between the race of the HH respondent and the average fuel efficiency of the vehicle owned is presented in Figure 3-5. There is no strong evidence that some ethnicities have a strong ‘’privilege’’ of purchasing and operating vehicles of higher fuel efficiency than others. For the national data, it seems that the “Asian only” category operates vehicles of higher fuel efficiency. On the other hand, for the local model, excluding the “Other” category, it seems that the “American Indian, Alaskan Native” category operates vehicles of higher fuel efficiency. Based on these observations, it shall be useful to identify if specific racial groups indicate particular behavior, via the development of dummy variables for particular ethnic groups.

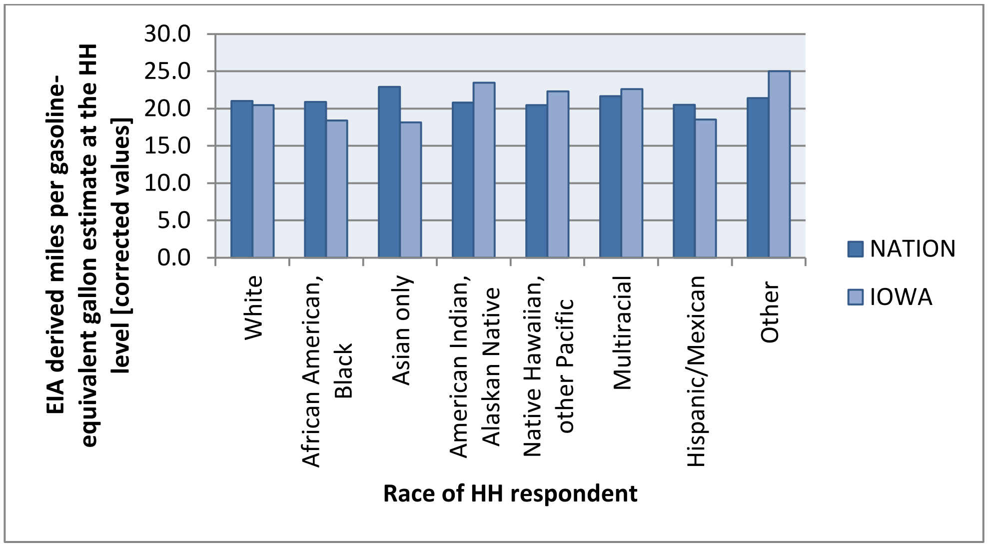

Figure 3-5: EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values] by race of HH respondent: Nation vs. Iowa

The relationship between HH income and the average fuel efficiency of the vehicle(s) owned by the HH is presented in Figure 3-6. The pattern is not very clear although generally it appears that, both at the national and local level, the higher the income, the higher the average vehicle fuel efficiency. Also, in a bid to identify whether income sub-groups may have a stronger explanatory power, it shall be useful to test for the following income sub-groups: LOW, MED, HIGH, VHIGH by setting the following lower and upper bounds:

Table 3-6: Average HH income sub-groups

Income Category	Lower bound	Upper bound
LOW	$2,500	$20,000
MED	$20,001	$40,000
HIGH	$40,001	$60,000
VHIGH	$60,001	$100,000

However, the fact that the pattern of Figure 3-6 is not clear may lead to the variable not being statistically significant in the average vehicle fuel efficiency at the HH level model.

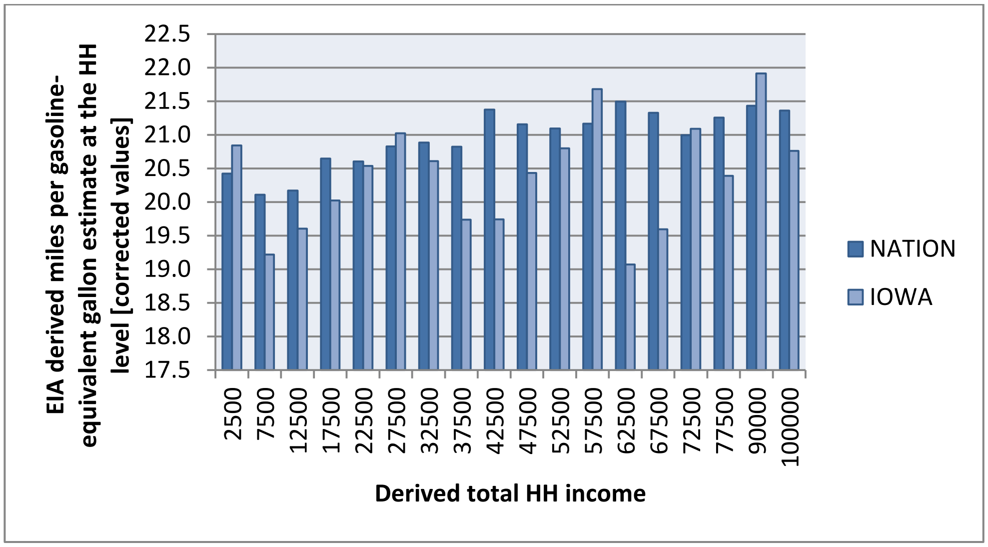

Figure 3-6: EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values] by derived total HH income: Nation vs. Iowa

Figure 3-7 does not provide clear evidence that some particular types of housing units are related to higher average vehicle fuel efficiency. Although this relationship is explored in a bid to treat the housing type as a proxy for average HH income, it still does not provide the author with a clear direction that can yield useful results for the final model specification.

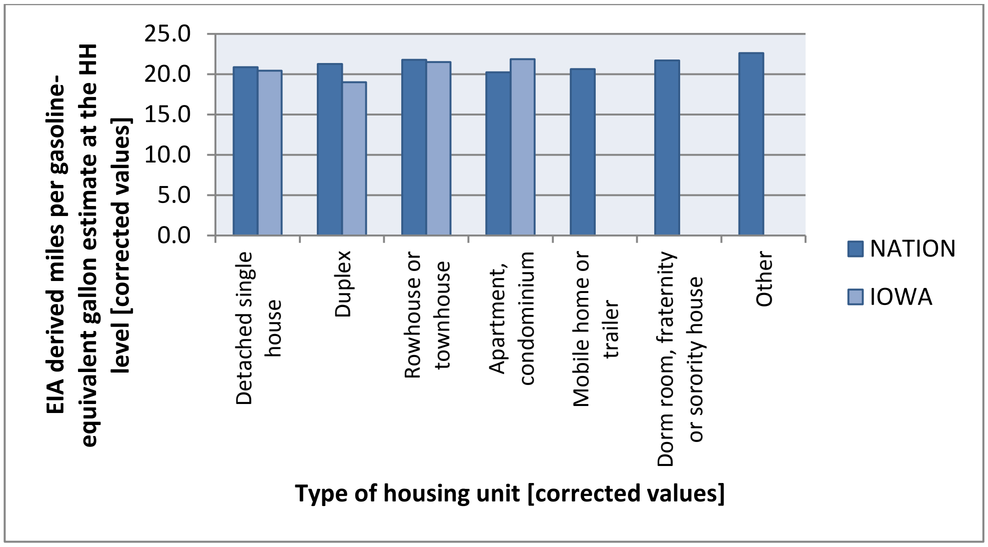

Figure 3-7: EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values] by type of housing unit [corrected values]: Nation vs. Iowa

Figure 3-8 provides evidence that as the number of workers per HH increases, so does the average fuel efficiency of the vehicle(s) owned at the HH level. This may be attributed to higher HH income, and may be related to the results of Figure 3-6. Interestingly enough, although for the national dataset the relationship is strictly increasing, in the case of Iowa there is an observed drop from 2 to 3 workers.

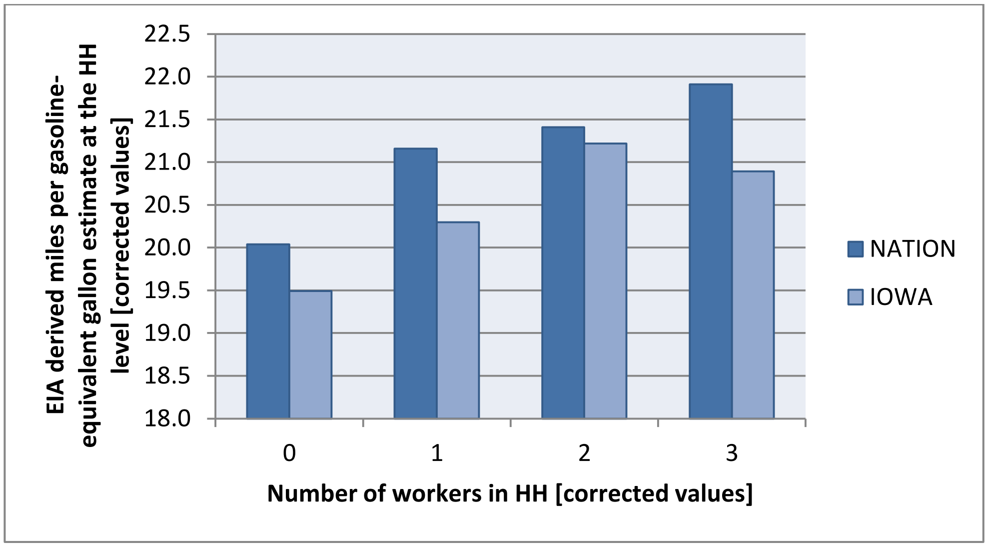

Figure 3-8: EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values] by number of workers in HH [corrected values]: Nation vs. Iowa

Figure 3-9 shows that, with regard to average vehicle fuel efficiency, there is a peak when the HH size is 3-4 people, but at the very limits of the range (1 and 6 people per HH), the average fuel efficiency drops. The local data follows the same pattern, but is slightly stronger than for the national data.

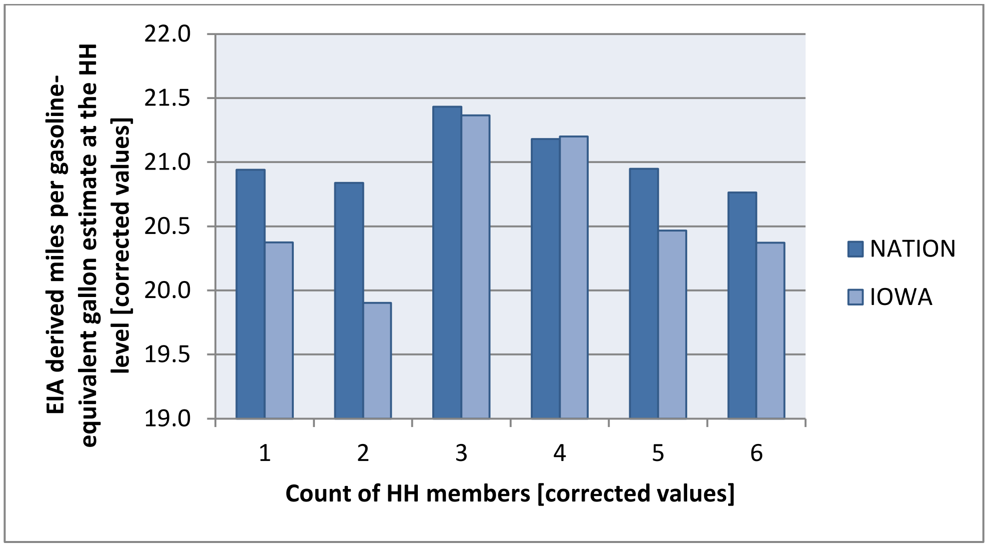

Figure 3-9: EIA derived miles per gasoline-equivalent gallon estimate at the HH level [corrected values] by count of HH members [corrected values]: Nation vs. Iowa

3.6.2. VMT fee – Y: Vehicle-miles traveled

For the VMT fee pricing option, BESTMILE is identified to be the most representative variable from the dataset. The goal is to identify which groups drive more, thus which groups will be mostly affected under a VMT fee alternative. Figure 3-10 shows the distribution of the best estimate of annual miles by home ownership status, both at the nationwide and local level. It is evident that people who own the housing unit they reside in tend to drive more at both levels of analysis. This result is justified if HH income is considered to be the factor which explains the home ownership status.

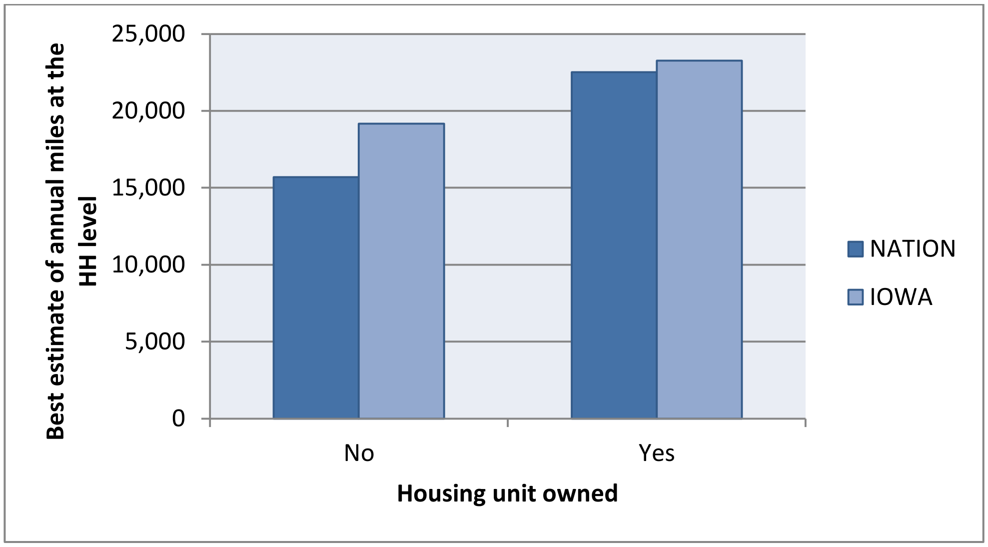

Figure 3-10: Best estimate of annual miles at the HH level by housing ownership status: Nation vs. Iowa

Similarly to the graph under the fuel efficiency section, access to rail reduces the number of trips by personal vehicle, as transit service is more competitive in terms of trip cost and delays in urban areas where congestion levels are problematic (Figure 3-11). It is noted, though, that for the local dataset, the RAIL variable is always equal to 0 (no heavy rail), thus the author cannot reach any useful conclusions regarding the local model, and the aforementioned variable shall not be included in the final local model specification (zero variable variation).

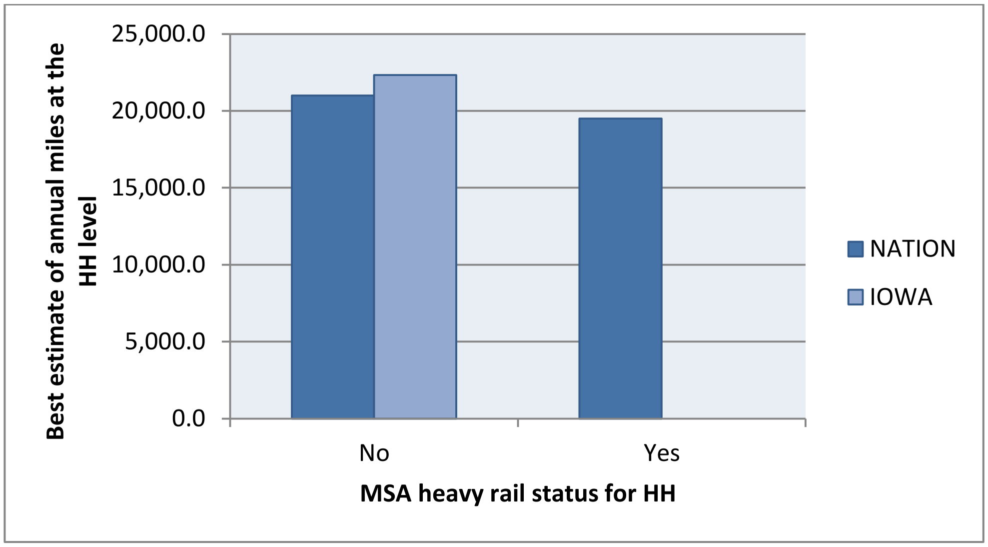

Figure 3-11: Best estimate of annual miles at the HH level by MSA heavy rail status for HH:
Nation vs. Iowa

The information provided in Figure 3-12 agrees with previous findings, in terms of fuel efficiency. For the national data, HHs with 2+ adults drive significantly more (almost twice as much) than their one-adult-HHs counterparts. This is expected as a higher number of residents generates a higher number of trips. Also, the presence of children affects the HH's VMT, since there is greater need to generate trips for various purposes. The peak, for the national data, is for the 2+ adults, youngest child 16-21 subcategory. This may be the case because the youngest child most probably is also a licensed driver and generates trips by himself/herself as well. For Iowa, the pattern described above is identical, though the absolute values of VMT are higher than the national figures for almost all the sub-groups of the HH life cycle classification.

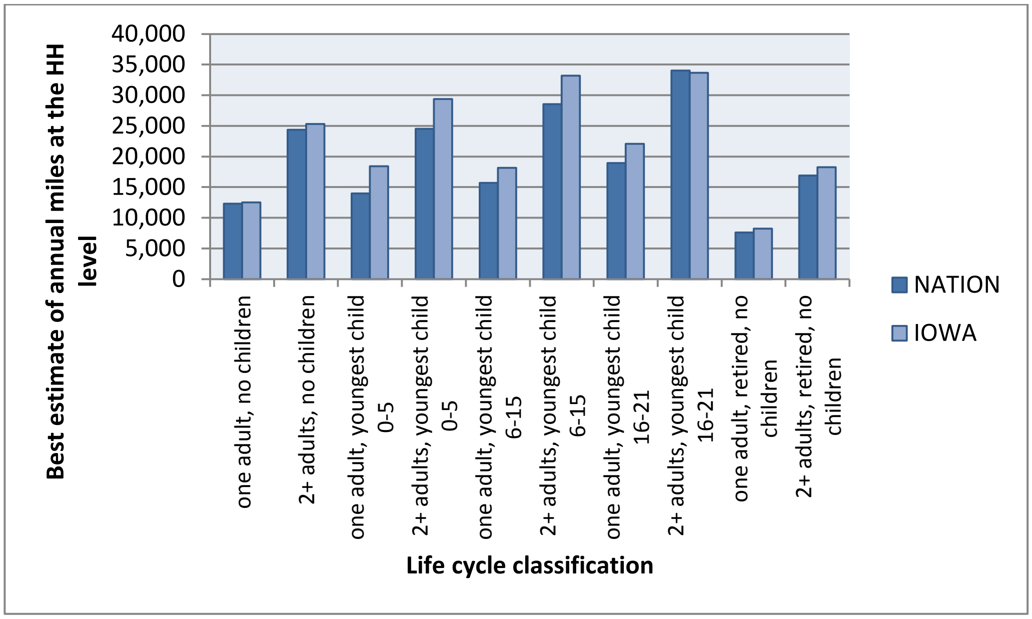

Figure 3-12: Best estimate of annual miles at the HH level by life cycle classification: Nation vs. Iowa

Figure 3-13 presents how BESTMILE changes by race group. At the national level, the racial sub-group that seems to drive more than the other sub-groups is the American Indian, Alaskan Native subgroup. Also, all the other subgroups have a quite similar VMT pattern. On the other hand, for the local data, the results have more peaks and lows. The racial sub-group which drives more is the Hispanic/Mexican, followed by the White subgroup. Even though this may seem peculiar due to the population composition of Iowa, it should be noted that the graph figures are not related to the associated frequency of each value. The only type of information that the author can reach is for comparison purposes.

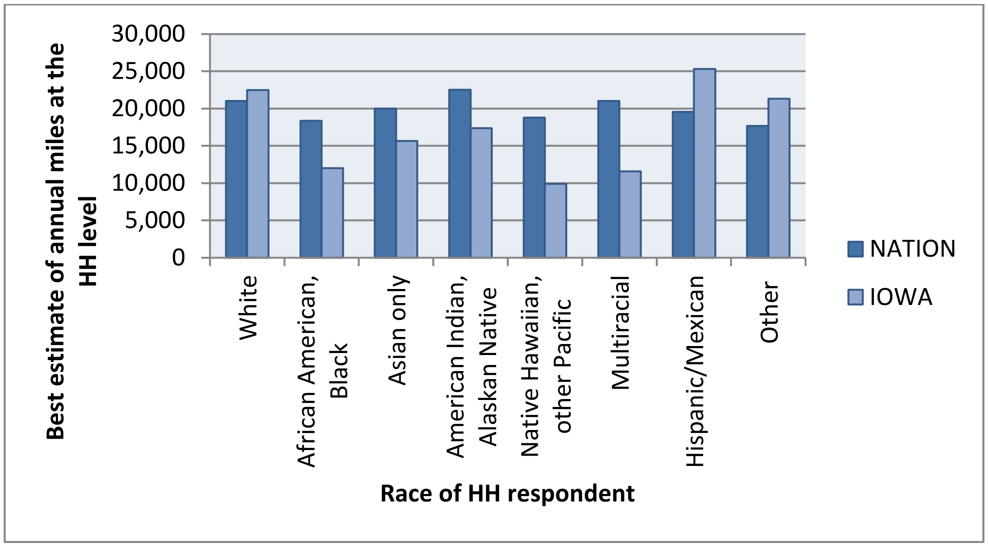

Figure 3-13: Best estimate of annual miles at the HH level by race of HH respondent: Nation vs. Iowa

In terms of average HH income, it seems that the higher the income, the more the HH residents drive (Figure 3-14). Here the increase is quite smooth, suggesting there is no justified reason why distinct income-subgroups should be tested separately in the model development of Chapter 4: Methodology and Results. The conclusions reached based on this graph are justified and supported by real-life practices where higher income groups drive more as the fuel-price burden is not as heavy as for lower-income groups. Also, higher VMT may be attributed to higher business activity that is usually related to higher incomes. Regarding the comparison between national and local data, the trend is similar across the two datasets, but it should be noted that for quite some income sub-groups, people in Iowa seem to drive more than their counterparts at the national level.

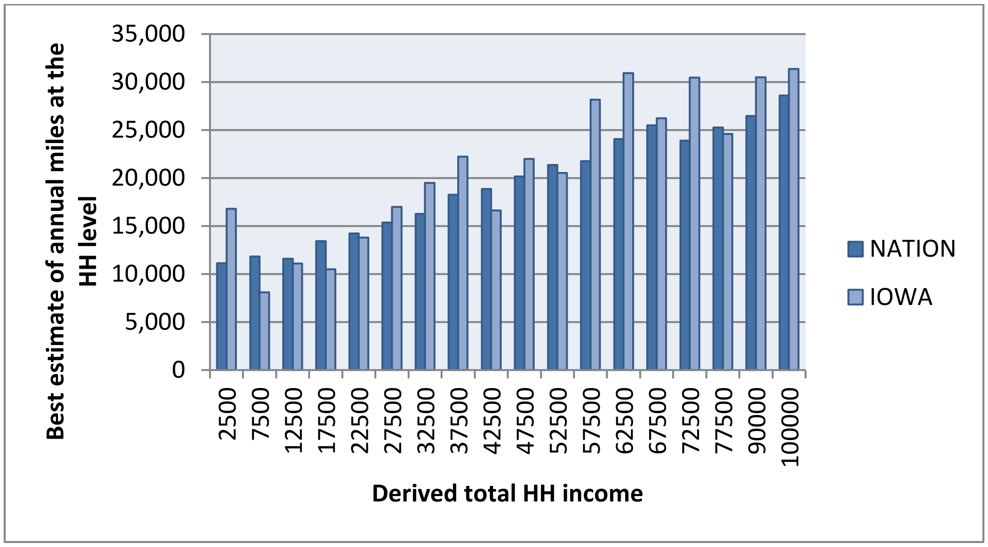

Figure 3-14: Best estimate of annual miles at the HH level by derived total HH income: Nation vs. Iowa

According to Figure 3-15, there are some observed differences in the average VMT at the HH level for different types of housing units. More precisely, residents of housing types which may be related to higher average HH income (such as detached single houses, or apartments/condominiums) seem to drive more than their counterparts in other housing types (rowhouses/townhouses, dorm rooms, etc.). This observation holds for both scales of analysis. It should be noted though that there is no valid information in the local dataset for the last three (3) sub-categories (mobile home or trailer, dorm room, fraternity or sorority house, and other).

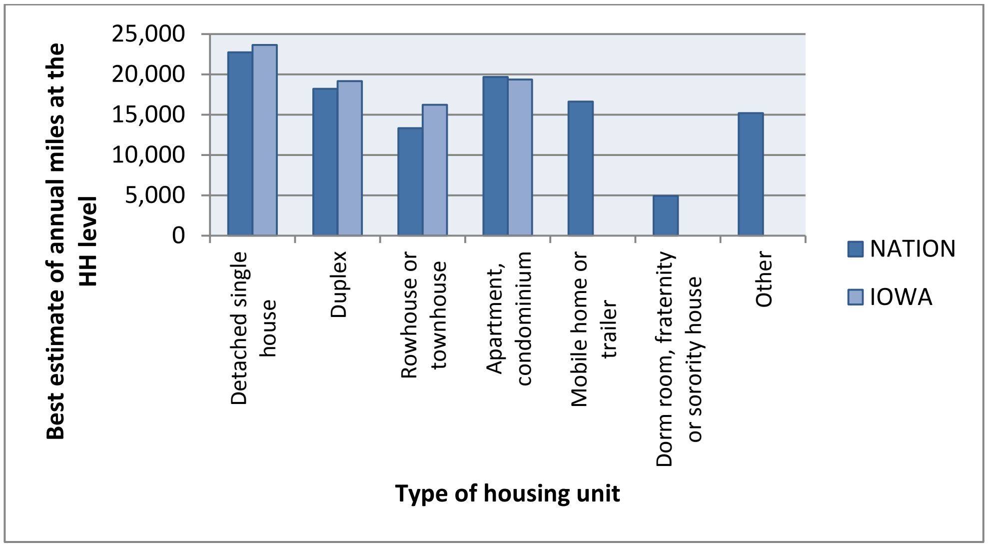

Figure 3-15: Best estimate of annual miles at the HH level by type of housing unit: Nation vs. Iowa

Similarly to the corresponding graphs under fuel efficiency, VMT increases as the number of workers per HH increases (Figure 3-16), due to the higher number of work-trips generated, which are more inelastic compared to other trip purposes. A similar trend is illustrated in Figure 3-17, where the larger the HH size, the higher the VMT. It should be noted, though, that according to Figure 3-17, the peak is for HHs of size around 4; when the HH size gets bigger than this “threshold”, the HH VMT slightly drops.

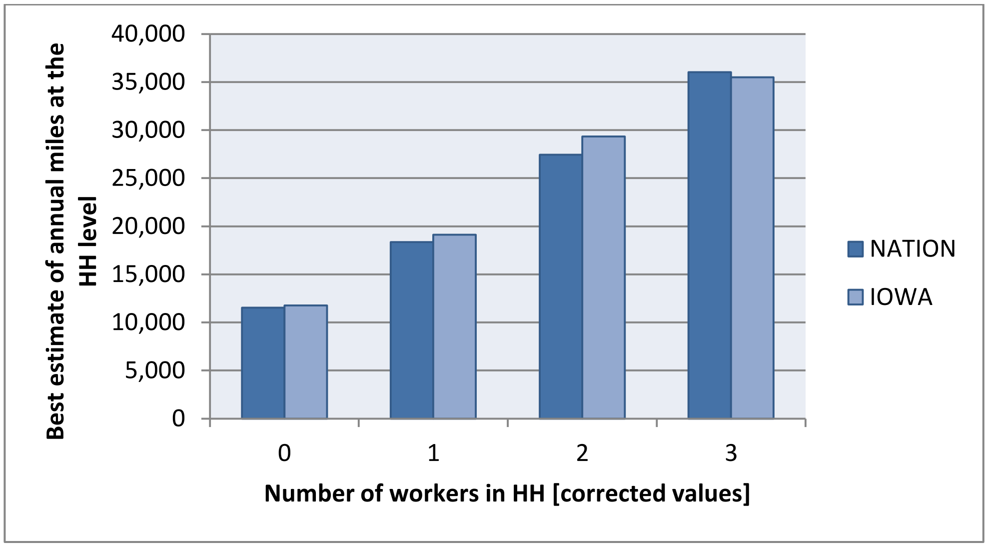

Figure 3-16: Best estimate of annual miles at the H level by number of workers in HH [corrected values]: Nation vs. Iowa

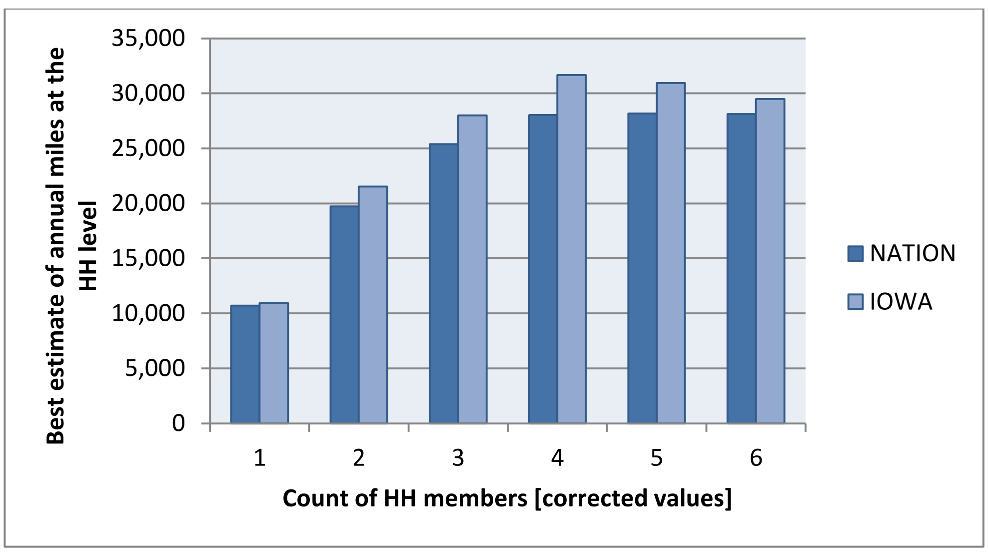

Figure 3-17: Best estimate of annual miles at the HH level by count of HH members [corrected values]: Nation vs. Iowa

3.7. Summary

The dataset used in the current analysis is Version 2.1 of the 2009 National Household Travel Survey (2009 NHTS). Upon merging, the final national dataset is at the HH level, and consists of 142,942 observations; the final dataset for the local model consists of 3,614 observations, for the state of Iowa. Additionally, in order to account for non-response, under-coverage, and multiple telephones in a HH, the data is weighted by the final HH weight, WTHHFIN. Upon merging the two distinct data files, it was deemed essential to remove some observations which are considered outliers. After removing the outliers, the final data still holds more than 96.13% of the original information.

Plotting of the variables is performed for both the national and the local datasets, in a bid to compare the figures and identify similarities and differences inferred by the graphs. For the fuel tax alternative, the EIA derived miles per gasoline-equivalent gallon estimate is considered to be a good representative of this pricing option. On the other hand, the best estimate of annual miles is the adequate choice for the VMT fee pricing option.

Regarding the average vehicle fuel efficiency at the HH level, respondents who own the home they reside in tend to operate vehicles of lower fuel efficiency, both at the national and the local level. Accessibility to rail leads to HHs with slightly higher vehicle fuel efficiency, although the RAIL variable displays no variation for the Iowa dataset. For the HH life cycle classification, the subcategory of one adult, with youngest child 0-5 years old operate vehicle of higher fuel efficiency than any other subgroup, for the national model. For the local data, this observation is valid for the 2+ adults, youngest child 6-15 subgroup. Additionally, regarding the HH race, the “Asian only” category operates vehicles of higher fuel efficiency at the national level; in Iowa, the “American Indian, Alaskan Native” category ranks higher. Income-wise, the pattern is not very clear although generally it appears that, both nationally and locally, the higher the income, the higher the average vehicle fuel efficiency. Finally, the higher the number of workers per HH, the higher the average vehicle fuel efficiency, whereas HHs of 3 to 4 members rank higher among other HHs.

Regarding the vehicle-miles traveled at the HH level, the conclusions inferred from the graphs are more straight-forward. In terms of the home ownership status, people who own the housing unit they reside in tend to drive more, at both levels of analysis. Access to rail reduces the number of personal-vehicle-trips, as transit service is more competitive in terms of trip cost and delays in urban areas where congestion levels are problematic; however, the RAIL variable displays no variation for the Iowa dataset. Also, regarding the HH life cycle classification, HHs with 2+ adults drive significantly more (almost twice as much) than their one-adult-HHs counterparts, at both levels. Additionally, at the national level, the American Indian, Alaskan Native subgroup drives more than the other sub-groups; at the local level, the racial sub-group which drives more is the Hispanic/Mexican, followed by the White subgroup. With regard to income, the higher the income, the more the HH residents drive, as the fuel-price burden is not as heavy as for lower-income groups. Furthermore, both nationally and locally, residents of housing types which may be related to higher average HH income seem to drive more than residents of other housing types. Finally, VMT increases as the number of workers per HH or as the HH size increases; this conclusion may be related to the income-specific conclusions reached above.

Chapter 4: Methodology and Results

4.1. Introduction

This chapter presents a methodology to identify the socioeconomic groups that are mostly affected under each pricing option. Identifying these groups is fundamental in roughly determining what groups are more likely to disproportionally shoulder the financial burden of the each option. To achieve this, the author identified a variable that is representative of the driver's fuel-related expenditures under each pricing option.

Under the current policy of the fuel tax per gallon, this thesis recommends the average fuel efficiency of the household fleet as a good proxy of the household's purchasing and travel behavior. More particularly, modeling the average fuel efficiency of the household fleet intends to identify which factors (socioeconomic, demographics, vehicle-related, or geographic) mostly affect this attribute as a vehicle choice in a household. The reason why this attribute is chosen as fair representative of the current pricing state is quite straight-forward; households which own vehicles of higher fuel efficiency tend to have lower fuel-related spending, whereas households which own less fuel efficient vehicles tend to consume more fuel, thus spend more. Similarly, under the alternative pricing option of the VMT fee, the average VMT by each household is selected as a suitable proxy for travel expenditures; identifying which social groups drive more will facilitate making inferences as to which social groups are most likely to carry the financial burden, once the VMT fee policy option is implemented.

Determining the relationship between fuel efficiency and VMT and the aforementioned independent variables is challenging, since these two models should be managed as a system, rather than as two distinct models, which do not interact with one another. More precisely, and as will be demonstrated in the following sections, the author has to deal with an interrelated system of equations where the dependent variable in one equation is the independent variable in another. Ignoring this pattern would most likely result in misspecified models and erroneous inferences. The kind of misspecification that the author tried to address via a simultaneous-equation model approach is endogeneity. The existence of endogeneity in a model suggests that a variable's variation (from the vector of independent variables) is caused by other exogenous or endogenous variables in the model (Washington, Karlaftis, & Mannering, 2011). This definition infers that an endogenous independent variable not only affects, but is also affected by the dependent variable.

4.2. Methodological approach

4.2.1. Three-stage least squares (3SLS)

Addressing the issue of model misspecification due to inherent endogeneity is accomplished via the application of the three-stage least squares estimator (3SLS). The 3SLS estimator is a system estimation method which "considers all of the parameter restrictions (caused by overidentification) in the entire equation system and accounts for possible contemporaneous (cross-equation) correlation of disturbance terms. Because system estimation approaches are able to utilize more information (parameter restrictions and contemporaneous correlation), they produce variance-covariance matrices that are at worst equal to, and in most cases smaller than, those produced by single-equation methods (resulting in lower standard errors and higher t-statistics for estimated model parameters)" (Washington, Karlaftis, & Mannering, 2011).

System-equation methods are typically preferred to single-equation methods because they account for restrictions in overidentified equations and contemporaneous (cross-equation) disturbance-term correlation (the correlation of disturbance terms across the equation system. In 3SLS, stage 1 is to get the two-stage least squares (2SLS) estimates of the model system. Stage 1 of 2SLS regresses each endogenous variable on all exogenous variables. Stage 2 of 2SLS uses regression-estimated values from stage 1 as instruments, and estimates each equation using OLS. In stage 2 of 3SLS, the 2SLS estimates are used to compute residuals from which cross-equation disturbance-term correlations are calculated. In stage 3, generalized least squares (GLS) is used to compute parameter estimates. Because of the additional information considered (contemporaneous correlation of disturbances), 3SLS produces more efficient parameter estimates than single-equation estimation methods. An exception is when there is no contemporaneous disturbance-term correlation. In this case, 2SLS and 3SLS parameter estimates are identical (Washington, Karlaftis, & Mannering, 2011).

4.2.2. Hypothesis testing – Analysis of Variance (ANOVA)

One of the research questions that this thesis attempts to address is whether a nationwide model for average vehicle fuel efficiency and vehicle miles traveled at the HH level can be successfully applied at a local level.

The analysis of variance (ANOVA) method can help specify whether the spatial level of analysis has a statistically significant effect on the observed average vehicle fuel efficiency and the VMT at the HH level.

According to Cobb (Cobb, 1998), “the evidence provided by the observed values is combined to form a special one-number summary, called an F-ratio. This one-number summary of the evidence summarizes a comparison: the average variability due to a particular factor of interest is compared with the average variability due to chance error.

$$F - ratio = Summary\ of\ the\ evidence =$$

$$= \frac{Average\ variability\ due\ to\ a\ particular\ factor}{Average\ variability\ due\ to\ chance\ error}$$

If this ratio is a lot bigger than 1, it means the average variability due to the difference in the factor of interest is a lot bigger than the average variability due to chance error. The analysis of variance table, or ANOVA table, is a handy table for keeping track of our work. The table has one row for each factor in the design and one column for each step in the analysis. In the case where the F-ratio is bigger than 1, we conclude that the factor effect is real, and we reject the null hypothesis:

$$H_o: there\ are\ no\ true\ differences\ due\ to\ the\ factor\ of\ interest''.$$

However, in order to identify if the detected differences due to the factor of interest are statistically significant, the F-ratio and the associated *p*-value is compared to the cut-off values, as those are defined at the level of significance of the study.

Note that, this testing procedure leads to inferences regarding solely the effect of the spatial level of analysis on the dependent variables considered in this thesis. It does not provide any kind of information regarding which set of variables affect the dependent variable(s) and to what extent. Thus, this hypothesis testing is only an introductory step in the comparison process between the national and the local model. Precise model development is required to determine the combination of factors which mostly affects the levels of the dependent variable(s), as well as the magnitude of this effect.

The local region of interest is the state of Iowa. The results of the ANOVA test are presented in Section 4.4.1, whereas the detailed model development for the local models is presented in Sections 4.4.3 and 4.4.4.

4.2.3. Asymptotic t-test of equality of individual coefficients between national and local model

As discussed in Section 4.2.2, the ANOVA methodological approach provides some preliminary information as to whether the spatial level of analysis has a statistically significant effect on the observed values on the average vehicle fuel efficiency and the vehicle-miles traveled at the HH level. However, the only source of comprehensive information on the relationship between dependent and independent variables comes from the model development.

According to Ben-Akiva & Lerman (1985), the asymptotic t-test of equality of individual coefficients is a powerful test which helps draw useful conclusions regarding the magnitude effect of the independent variables on the dependent ones, and how these compare across two models.

In generic terms, the asymptotic t-test of equality of individual coefficients between model 1 and model 2 is defined as follows:

$$\frac{\widehat{\beta_k^1} - \widehat{\beta_k^2}}{\sqrt{var\left(\widehat{\beta_k^1}\right) + var\left(\widehat{\beta_k^2}\right) - 2cov\left(\widehat{\beta_k^1}, \widehat{\beta_k^2}\right)}} \sim t^*_{df,\alpha}$$

The asymptotic t-test is conducted for each (common) variable. The computed t-value is compared to the t-table value for the considered degrees of freedom and level of significance($t^*_{df,\alpha}$), under the following null hypothesis:

Ho: $there\ is\ no\ statistically\ significant\ difference\ between\ the\ coefficients\ of$ $model\ 1\ and\ model\ 2$ (Ben-Akiva & Lerman, 1985).

4.3. Analysis Results: National Model

The current section presents the model results of the 3SLS estimation, for the national model. First, the 3SLS model specification is presented, followed by the results of the model estimation. The results are discussed in terms of the goodness of fit (GOF) of the model, the parameter estimation and the associated elasticities (with respect to fuel efficiency and VMT), as well as the computed elasticities.

4.3.1. 3SLS Model Specification

The two models are managed as a system with the following combination of variables (please refer to Table 3-2: Weighted descriptive statistics – National Data for more information on the variables):

$$Leffc =$$

$$= f(Lvmt, Large, Wrktosiz, Vehownm2, Vehage, Homeown, Rail, Tax, Urb, Nosub, Oneadult, Land, Hybrid)$$

$$Lvmt =$$

$$= f(Leffc, Nosub, Drvrcnt2, Lincome, Wrktosiz, Large, Urb, Lmsasiz, Tax, Land, Vehownm2, Rail)$$

The endogeneity issue of the model due to the variables LEFFC and LVMT is accounted for via the 3SLS estimator, where the vector of the instrumental (exogenous) variables includes:

$$Instrumental\ Variables =$$
$$= \{Large, Wrktosiz, Vehownm2, Vehage, Homeown, Rail, Tax, Urb, Nosub, Oneadult, Land, Hybrid, Drvrcnt2, Lincome, Lmsasiz\}$$

It is pointed out that in order to obtain valid results - representative of the dataset utilized - the appropriate household weights (WTHHFIN) have been applied, as discussed in Chapter 3: Data Description. The final parameter estimations of the previewed model-system are presented in Sections 4.3.2 and 4.3.3. Even though the model estimation is conducted system-wise, the results are presented separately, in order to facilitate the discussion of each pricing option, in the context of the current thesis.

4.3.2. Fuel tax pricing option: 3SLS model results

The current section presents the estimations for the fuel efficiency model. Table 4-1 presents the parameter coefficients for each independent variable, as those have been computed via the 3SLS estimator. For ease of discussion, the independent variables have been categorized into three sub-groups: vehicle, geographic, and socioeconomic characteristics. Each parameter estimation is presented along with its associated *t*-statistic and *p*-value, indicating its level of statistical significance. Additionally, Table 4-1 displays the number of variables used for fitting the suggested model, the GOF measures of R^2 and adjusted R^2, as well as the Durbin-Watson statistic, as a measure of presence of spatial autocorrelation in the model.

Table 4-1: Model results – Y: Average vehicle fuel efficiency at the HH level - Nation

Variable Mnemonic	**Coefficient**	***t*-statistic**	***p*-value**
Constant	2.139	82.24	0.0000
Vehicle characteristics			
HYBRID	0.014	6.13	0.0000
NOSUB	0.105	4.41	0.0000
VEHOWNM2	-0.0002	-10.03	0.0000
VEHAGE	-0.004	-31.46	0.0000
LARGE	-0.182	-150.56	0.0000
LVMT	0.103	39.88	0.0000
Geographic characteristics			
LAND	-0.009	-6.08	0.0000
URB	0.018	14.77	0.0000
RAIL	0.012	10.53	0.0000
TAX	0.036	7.96	0.0000
Socio-economic characteristics			
WRKTOSIZ	0.036	21.83	0.0000
HOMEOWN	-0.013	-13.17	0.0000
ONEADULT	0.012	8.82	0.0000
No. of observations		116,045	
R-squared		0.372	
Adjusted R-squared		0.372	
Durbin-Watson		2.016	

** All variables are statistically significant at a 99.99% level of significance (α = 0.0001)*

Discussion of results

The selection of the variables included in the model specification was made based on the level of correlation of each independent variable with the dependent variable of the particular model (LEFFC) (Appendix 2A: Correlation Matrix – Nation). Those variables which were highly correlated (in absolute value) to the dependent variable were selected first, and the model building continued in descending order. This selection did not include variables which are related to the dependent variable via a straight-forward mathematical formula. Some variables of the dataset at derived from the combination of other variables, via a mathematical expression. In this case, only the original or the derived variables were used at a time, to avoid correlation issues in the model specification. All independent variables in the final model have the expected sign, according to the correlation matrix.

Goodness of fit

The fuel efficiency model displays a fair fit to the data, indicating that 37.2% of the inherent data variability can be explained by the proposed model. The nature of the dataset used – nationwide travel survey data – in conjunction with its inherent high variability justifies the level of model fit; and could suggest that the model explains little of a lot of variability. Additionally, the value of the Durbin-Watson statistics at 2.016 suggests that the model does not suffer from spatial autocorrelation.

Parameter estimation and elasticities

The current section discusses the magnitude of the estimated coefficients, as those are presented in Table 4-1. The discussion focuses on the computed elasticities with respect to the natural logarithm of fuel efficiency (LEFFC), as those are derived from the estimated coefficients.

- **Vehicle Characteristics**

As it has been well expected, the vehicle attributes define its fuel efficiency, thus the model includes a good number of vehicle attributes as predictors of the average vehicle fuel efficiency at the HH level.

Hybrid vehicles have higher fuel efficiency than conventional vehicles or vehicles which consume other types of fuel (most usually, conventional fuel such as gasoline or diesel). HHs with hybrid vehicles or vehicles using alternative fuel have 0.014% higher average vehicle fuel efficiency than their counterparts operating conventional vehicles. This may suggest, although it is not implied by the current model, that such households are more likely to have a higher income per capita than HHs which own conventional vehicles, as their purchase price is usually higher, and acts as a repellent for the low- or medium-income HHs. The number of vehicles per household, as this is captured via the dummy NOSUB factor, affects positively the average vehicle fuel efficiency of the household, indicating that HHs which own more than one vehicle, tend to choose vehicles of approximately 0.105% higher fuel efficiency. This also may be related with the income variable, suggesting that HHs who own more vehicles are located high in the income rating list, whereas HHs with only one vehicle may be located lower in the same list, and attempt to reduce their fuel-related expenditures via the purchase of a fuel efficient vehicle. Additionally, households which own ‘’large’’ vehicles, i.e. pick-up trucks and SUVs, have approximately 0.182% lower average vehicle fuel efficiency than other vehicle types, such as automobiles. The information provided by this variable is also related to the fact that ownership of such vehicle type is most common in less urbanized areas.

The vehicle ownership period has a negative effect on the fuel efficiency under which the vehicle is performing; this suggests that HHs who own their vehicles for a long time are most likely to drive less fuel efficient vehicles. The information provided by this variable is related to the HH income and the vehicle age. HHs which own their vehicles for a longer time may be listed lower in the income scale (lack of economic ability to renew their vehicle), and also are more likely to own a vehicle which does not meet the current standards for fuel efficiency, as it was manufactured in a time period when those standards were lower. Moreover, the endogenous variable of VMT shows that HHs which have more travel-related activities choose to operate fuel efficient vehicles, leading to possibly lower fuel consumption. In other words, HHs which generate a large amount of trips are more likely to invest on a fuel efficient vehicle, in order to balance the fuel expenditures which results from their high travel activity. In elasticity terms, 1% in the HH's annual VMT leads to an approximate 0.103% higher fuel efficiency of the vehicles operated by the HH.

- **Geographic Characteristics**

The location-specific attributes affect significantly the model development. Households located in urban areas, where the rail alternative is present, tend to choose vehicles of 0.018% higher fuel efficiency than in rural areas, and 0.012% than in areas where there is no mode alternative. In areas where the taxation per fuel gallon including federal and state tax) is higher, HHs tend to drive vehicles of higher fuel efficiency. This finding is reasonably supported by the fact that HHs try to reduce their fuel-related expenditures, by driving vehicles which consume less for the same distance traveled. Also, HHs located in residential areas operate vehicles of 0.009% lower fuel efficiency than HHs located in areas with mixed land use.

- **Socioeconomic Characteristics**

  The impact of the socioeconomic attributes on the average fuel efficiency of the vehicles owned by the household is captured through the variables WRKTOSIZ, ONEADULT, HOMEOWN. It is inferred that HHs which consist of only one adult versus HHs with more than one adult, tend to drive vehicles of 0.012% higher mpg. Similarly to previously discussed variables, WRKTOSIZ also serves as a surrogate for income suggesting that higher WKRTOSIZ values may suggest higher HH income. HHs with higher workers to HH size ratio tend to drive more fuel efficient vehicles as well (0.036% higher fuel efficiency). On the other hand, HHs which own their housing unit are less likely to drive fuel efficient vehicles than their counterparts who rent it (0.013% lower fuel efficiency).

It is noted that the factor of income did not have the expected sign and therefore was not included in the fuel efficiency model specification; however, other variables included in the model, such as HOMEOWN, WRKTOSIZ, or NOSUB, may indirectly capture the effect of income on the average vehicle fuel efficiency at the HH level. It is also noted that, the assumption that sub-groups who purchase and operate fuel efficient vehicles carry less of the fuel tax burden, should be also supplemented by the assumption that this purchase is ‘’induced’’ by the prevailing conditions, and should also be considered part of their total expenditures under the pricing option of fuel taxation.

4.3.3. VMT fee pricing option: 3SLS model results

The current section presents the estimations for the VMT model. Table 4-2 presents the parameter coefficients for each independent variable, as those have been computed via the 3SLS estimator. For ease of discussion, the independent variables have been categorized into three sub-groups; vehicle characteristics, geographic characteristics, and socioeconomic characteristics. Each parameter estimation is presented along with its associated t-statistic and p-value, indicating its level of statistical significance. Additionally, Table 4-2 displays the number of variables used for fitting the suggested model, the GOF measures of R^2 and adjusted R^2, as well as the Durbin-Watson statistic, as a measure of presence of spatial autocorrelation in the model.

Table 4-2: Model results – Y: VMT at the HH level - Nation

Variable Mnemonic	**Coefficient**	***t*-statistic**	***p*-value**
Constant	0.835	3.25	0.0011**
Vehicle characteristics			
LEFFC	2.291	25.59	0.0000*
NOSUB	-0.453	-68.63	0.0000*
LARGE	0.552	36.54	0.0000*
VEHOWNM2	-0.002	-18.24	0.0000*
Geographic characteristics			
URB	-0.141	-27.87	0.0000*
LMSASIZ	-0.004	-11.25	0.0000*
TAX	-0.187	-10.01	0.0000*
LAND	0.020	3.32	0.0009**
RAIL	-0.051	-10.24	0.0000*
Socio-economic characteristics			
DRVRCNT2	0.218	53.56	0.0000*
LINCOME	0.140	42.87	0.0000*
WRKTOSIZ	0.134	16.31	0.0000*
No. of observations		116,045	
R-squared		0.535	

Variable Mnemonic	Coefficient	*t*-statistic	*p*-value
Adjusted R-squared		0.535	
Durbin-Watson		1.988	

** Variable is statistically significant at a 99.99% level of significance (α = 0.0001)*
*** Variable is statistically significant at a 99% level of significance (α = 0.01)*

Discussion of results

The selection of the variables included in the model specification was made based on the level of correlation of each independent variable with the dependent variables of the particular model (LVMT) (Appendix 2B: Correlation Matrix – Iowa); those variables which are highly correlated (in absolute value) to the dependent variable were selected first, and the model building continued in descending order. All independent variables in the final model have the expected sign, according to the correlation matrix.

Goodness of fit

The VMT model displays a better fit to the data, indicating that 53.5% of the inherent data variability can be explained by the proposed model. Similarly to the explanation under the fuel efficiency model, it has quite good explanatory power for such a large dataset, following the rule of thumb that the higher the inherent variability of the dataset, the lower the model fit. Additionally, the Durbin-Watson statistic is equal to 1.988, suggesting there are no spatial autocorrelation issues in the model.

Parameter estimation and elasticities

The current section discusses the magnitude of the estimated coefficients, as those are presented in Table 4-2. The discussion focuses on the computed elasticities with respect to VMT, as those are derived from the estimated coefficients.

- **Vehicle Characteristics**

  As expected, the average vehicle fuel efficiency is a statistically significant factor which positively affects the vehicle miles traveled at the household level; the higher the average vehicle fuel efficiency of the household, the higher the number of vehicle miles traveled at this level of analysis. More precisely, 1% increase in the average vehicle fuel efficiency may lead to 2.291% increase in the VMT at the HH level, making it the only variable with elastic performance in the model. This finding is supported by the idea that fuel efficient vehicles produce higher savings for the same distance traveled, thus households that own such vehicles are induced to travel more. The number of vehicles per household, as this is captured via the dummy NOSUB factor, affects negatively the VMT of the household; lower vehicle ownership levels lead to limited travel-related activities of approximately 0.453%. Additionally, households which own ‘’large’’ vehicles, i.e. pick-up trucks and SUVs, tend to have 0.552% higher VMT compared to other vehicle types, such as automobiles, most probably because these vehicles are owned by HHs which are located in less urbanized areas where activities are decentralized and there is no high availability of alternative modes. Finally, the time period (in months) along which the vehicle has been owned by the household interacts positively with the VMT, suggesting that the longer the ownership period, the higher the miles traveled at the household level. This may be attributed to the concept of induced travel that goes along with new vehicle purchases.

- **Geographic Characteristics**

The geographic attributes seem to also affect significantly the vehicle miles traveled. Households located in urban areas are found to drive 0.141% less than their rural counterparts, due to decentralized activities (commercial and other) and due to significantly lower housing density. In the same context, households located in bigger MSA (metropolitan statistical areas) seem to drive less. 1% increase in the population size of the MSA leads to 0.004% lower VMT. This is similar to the explanation for the urban vs. rural case, where urban sprawl affects trip generation, trip distribution, and trip duration. Additionally, the level of taxation per fuel gallon (including both federal and state tax) has a negative impact on how much people drive, as it reflects the level of the vehicle operating cost. 1 dollar increase in the fuel tax leads to 0.187% reduction in the HH’s VMT. Also, households which are located in residential areas are found to drive 0.020% more than households located in business districts, most probably due to homogeneous nature of the area, which mainly consists of residential units, and does not leave room for commercial and other types of activities to develop. Finally, the presence of rail in the area reduces the VMT of the household by 0.051%. Even though this variable captures the presence of heavy rail, it can still be argued that it serves as a surrogate for public transit, and is used as a mode alternative mainly for long-distance trips.

- **Socioeconomic Characteristics**

  Regarding the socio-economic characteristics of the household, higher income households tend to drive more than their lower-income counterparts, as expected. 1% increase in the HH income leads to a 0.14% increase in the VMT of the HH. This result agrees with the magnitude effect of WKRTOSIZ variable, i.e. the ratio of number of workers in the HH over the whole HH size. The larger this ratio, the higher the travel-related activities, as more work-trips are required and generated. (elasticity of 0.134%), suggesting that income plays an important role in the travel behavior of the HH. Finally, the higher the number of licensed drivers in the HH, the more likely they are to drive more. An additional driver increases the HH VMT by 0.218%.

4.4. Analysis Results: Local model for Iowa

The second objective of this thesis is to identify whether there is a statistically supported need for a local model in order to explain and predict the average vehicle fuel efficiency and the VMT at the HH level for a local region. The development of both a nationwide and a local model is the goal of this thesis, in a bid to identify if there is a need for such a distinction, if they share the same or similar combination of influential factors, and whether all those differences are supported statistically.

The region selected for this analysis is the state of Iowa. The comparison analysis consists of three distinct steps. First, the author studies the effect of region (Nation vs. Iowa) on the dependent variables of interest, via the ANOVA analysis presented in Sections 4.4.1.1 and 4.4.1.2 for the average vehicle fuel efficiency and VMT at the HH level respectively.

4.4.1. Comparison of the nationwide model and the Iowa model

4.4.1.1. ANOVA results: Fuel efficiency model

The current section studies whether the effect of region is statistically significant on the average vehicle fuel efficiency at the HH level. More precisely, it is of interest to identify whether there are detectable differences in the average vehicle fuel efficiency at the HH level that can be attributed to the level of regional analysis (national vs. local). In statistical terms, the null hypothesis is formed as follows:

$$H_o: there\ are\ no\ true\ differences\ in\ the\ average\ vehicle\ fuel\ efficency\ at\ the\ HH\ level\ due\ to\ the\ level\ of\ regional\ analysis\ (national\ vs.local)$$

The following tables present the detailed ANOVA performed for the average vehicle fuel efficiency of the nationwide model and the local model (Iowa).

Table 4-3: ANOVA – Effect of region on average vehicle fuel efficiency at the HH level

Effect Tests					
Source	**Nparm**	**DF**	**Sum of Squares**	**F-ratio**	**Prob>F**
Region	1	1	0.71979848	0.7866	0.3751

The ANOVA analysis for the effect of the level of regional analysis indicates that there are no statistically significant differences in the average vehicle fuel efficiency that can be attributed to the effect of the level of spatial analysis (nationwide vs. local) at 90% level of significance (p-value = 0.3751>0.10).

4.4.1.2. ANOVA results: VMT model

The current section studies whether the effect of region is statistically significant on the vehicle miles traveled at the HH level. More precisely, we are interested to identify whether there are detectable differences in the vehicle miles traveled at the HH level that can be attributed to the level of regional analysis (national vs. local). In statistical terms, the null hypothesis is formed as follows:

$$H_o: there\ are\ no\ true\ differences\ in\ the\ vehicle\ miles\ traveled\ at\ the\ HH\ level\ due\ to\ the\ level\ of\ regional\ analysis\ (national\ vs.local)$$

The following tables present the detailed ANOVA performed for the vehicle miles traveled at the HH level of the nationwide model and the local model (Iowa).

Table 4-4: ANOVA – Effect of region on VMT at the HH level

Effect Tests					
Source	**Nparm**	**DF**	**Sum of Squares**	**F-ratio**	**Prob>F**
Region	1	1	0.10236362	2.662	0.1028

The ANOVA analysis for the effect of the level of regional analysis indicates that there are statistically significant differences in the vehicle miles traveled that can be attributed to the effect of the level of spatial analysis (nationwide vs. local) (level of regional analysis is marginally significant at 90% level of significance - *p*-value = 0.1028 ≈ 0.10). Note that, the effect of spatial analysis on VMT is stronger than its effect on average vehicle fuel efficiency.

Discussion of results

The first ANOVA test suggests that no statistically significant differences have been detected in the observed values of the average vehicle fuel efficiency that are attributed to the spatial effect (national vs. local model); on the other, the spatial effect has a marginally statistically significant effect on the VMT at the household level. However, we still need to identify whether the same factors affect the dependent variables (vector of independent variables), towards what direction (coefficient sign) and to what extent (coefficient magnitude).

4.4.2. 3SLS Model Specification

Similarly to the national 3SLS model specification, the two models are managed as a system with the following combination of variables (please refer to Table 3-4: Weighted descriptive statistics: Local data (Iowa) for more information on the variables):

$$Leffc = f(Lvmt, Hybrid, Vehownm2, Vegage, Large, Wrktosiz, Lmsasiz)$$

$$Lvmt = f(Leffc, Nosub, Large, Urb, Lmsasiz, Lincome, Wrktosiz, Oneadult)$$

The endogeneity issue of the model due to the variables LEFFC and LVMT is accounted for via the 3SLS estimator, where the vector of the instrumental (exogenous) variables includes:

$$Instrumental\ Variables = \{Large, Wrktosiz, Vehownm2, Vehage,$$
$$Urb, Nosub, Oneadult, Hybrid, Lincome, Lmsasiz\}$$

It is pointed out that in order to obtain valid results - representative of the dataset utilized - the appropriate household weights (WTHHFIN) have been applied. The final parameter estimations of the previewed model-system are presented in Sections 4.4.3 and 4.4.4. Even though the model estimation was conducted system-wise, the results are presented separately, in order to facilitate the discussion of each pricing option, in the context of this thesis.

4.4.3. Fuel tax pricing option: 3SLS model results

Table 4-5: Model results – Y: Average vehicle fuel efficiency at the HH level - Iowa

Variable Mnemonic	Coefficient	*t*-statistic	*p*-value
Constant	2.447	36.058	0.0000*
Vehicle Characteristics			
HYBRID	-0.037	-3.546	0.0004**
VEHOWNM2	-0.0004	-5.114	0.0000*
VEHAGE	-0.004	-5.243	0.0000*
LARGE	-0.134	-18.628	0.0000*
LVMT	0.070	9.718	0.0000*
Geographic characteristics			
LMSASIZ	0.002	3.528	0.0004**
Socio-economic characteristics			
WRKTOSIZ	0.034	3.922	0.0001**
No. of observations		2,994	
R-squared		0.279	
Adjusted R-squared		0.277	

* *Variable is statistically significant at a 99.99% level of significance (α = 0.0001)*

** *Variable is statistically significant at a 99% level of significance (α = 0.01)*

Discussion of results

Goodness of fit

The fuel efficiency model displays a fair fit to the data, indicating that 27.7% of the inherent data variability can be explained by the proposed model. The nature of the dataset used – nationwide information – in conjunction with its inherent high variability justifies the level of model fit; *the suggested model explains little of a lot of variability.* Additionally, using such a widely conducted national survey to explain the variability in such a detailed level (local model) does not provide us with the detailed information necessary to develop a model that fits the data better.

Parameter estimation and elasticities

The current section discusses the magnitude of the estimated coefficients, as those are presented in Table 4-5, while taking into account the computed elasticities with respect to VMT.

- **Vehicle Characteristics**

  As it has been well expected, the vehicle attributes define its fuel efficiency, thus the model includes a good number of vehicle attributes as predictors of the average vehicle fuel efficiency at the HH level. Contrary to the results of the national model, HHs which own hybrid vehicles seem to have lower average fuel efficiency (0.037%) that HHs which own only conventional vehicles or vehicles which consume other types of fuel (most usually, conventional fuel such as gasoline or diesel), a finding which agrees with the expected sign from the correlation matrix for the Iowa data. This may be true in the event that HHs which own a hybrid vehicle also own vehicles of lower fuel efficiency, such as pick-up trucks or SUVs, thus the average fuel efficiency at the HH level is skewed towards the lower values, based on level of usage. Additionally, households which own ‘’large’’ vehicles, i.e. pick-up trucks and SUVs, have lower (0.134%) average vehicle fuel efficiency than other vehicle types, such as automobiles. Similarly, HHs with older vehicles have lower average vehicle fuel efficiency than their counterparts with newer vehicles. More precisely, an additional year of vehicle life reduces its fuel efficiency by 0.004%. Additionally, the time period (in months) along which the vehicle has been owned by the household interacts negatively with the average vehicle fuel efficiency at the HH level, suggesting that the longer the ownership period, lower the fuel efficiency of the vehicle (this variable contains similar information to the VEHAGE variable). In elasticity terms, one additional moth in the ownership period decreases the average vehicle fuel efficiency at the HH level by 0.0004%. Finally, HHs which drive more as a result of high travel-related activity tend to choose vehicles of higher fuel efficiency. In elasticity terms, a 1% increase in the HH VMT leads to 0.07% increase in the average vehicle fuel efficiency at the HH level.

- **Geographic Characteristics**

  For the Iowa-specific model, geographic characteristics do not seem to have a statistically significant effect on the average vehicle fuel efficiency at the household level. This is reasonable given the fact that the geographic-specific variables included in the 2009 NHTS contain information which is too generic/ aggregated, thus displays no variation along the Iowa observations. The only location-specific variable included in the model specification is LMSASIZ, indicating that HHs located in larger MSAs (metropolitan statistical areas) tend to drive vehicles of higher fuel efficiency. 1% increase in the MSA size leads to 0.002% increase in the average vehicle fuel efficiency at the HH level.

- **Socioeconomic Characteristics**

  The impact of the socioeconomic attributes on the average fuel efficiency of the vehicles owned by the household is captured through the variable WRKTOSIZ. It is inferred that HHs with higher workers to HH size ratio - which also serves as a proxy for HH income - tend to drive more fuel efficient vehicles (elasticity of 0.034%)

4.4.4. VMT fee pricing option: 3SLS model results

Table 4-6: Model results – Y: VMT at the HH level – Iowa

Variable Mnemonic	**Coefficient**	***t*-statistic**	***p*-value**
Constant	-0.833	-0.78	0.4347
Vehicle Characteristics			
LEFFC	2.789	7.42	0.0000*
NOSUB	-0.499	-12.37	0.0000*
LARGE	0.498	9.35	0.0000*
Geographic characteristics			
URB	-0.155	-6.68	0.0000*
LMSASIZ	-0.009	-4.22	0.0000*
Socio-economic characteristics			
LINCOME	0.185	8.95	0.0000*
WRKTOSIZ	0.170	4.40	0.0000*
ONEADULT	-0.209	-6.29	0.0000*
No. of observations		2,985	
R-squared		0.480	
Adjusted R-squared		0.479	

** Variable is statistically significant at a 99.99% level of significance ($\alpha = 0.0001$)*

Discussion of results

Goodness of fit

The VMT model displays a better fit to the data than the model for the average vehicle fuel efficiency, indicating that 47.9% of the inherent data variability can be explained by the proposed model. Similarly to the explanation under the fuel efficiency model, the model has quite good explanatory power. It should be pointed out that, even though the Iowa dataset is quite small, the fit is not very high due to the non-detailed information provided by the variables of the dataset. It is most likely that for such a localized model we would need variables which would describe the current travel behavior at a greater/ more detailed level of analysis.

Parameter estimation and elasticities

The current section discusses the magnitude of the estimated coefficients, as those are presented in Table 4-7, while taking into account the computed elasticities with respect to VMT.

- **Vehicle Characteristics**

  As expected, the average vehicle fuel efficiency is a statistically significant factor which positively affects the vehicle miles traveled at the household level; the higher the average vehicle fuel efficiency of the household, the higher the number of vehicle miles traveled at this level of analysis. In elasticity terms, 1% increase in VMT leads to a 2.789% increase in the average fuel efficiency of the vehicles that the HH chooses to operate, making it the only variable with elastic performance in the model. This finding is supported by the idea that fuel efficient vehicles produce higher savings for the same distance traveled, thus households that own such vehicles are induced to travel more. The number of vehicles per household, as this is captured via the dummy NOSUB factor, affects negatively the VMT of the household; lower vehicle ownership levels lead to limited travel-related activities (elasticity of 0.499%). Additionally, households which own ‘’large’’ vehicles, i.e. pick-up trucks and SUVs, tend to drive more (0.498%) compared to other vehicle types, such as automobiles.

- **Geographic Characteristics**

  The geographic attributes seem to also affect significantly the vehicle miles traveled. Households located in urban areas are found to drive less than their rural counterparts (0.155%), possibly because of the availability of other alternatives, such as public transit. In the same context, households located in bigger MSA (metropolitan statistical areas) seem to drive less. In elasticity terms, 1% increase in the MSA size leads to 0.009% decrease in the VMT at the HH level.

- **Socioeconomic Characteristics**

  Higher-income households tend to drive more than their lower-income counterparts, as expected. In elasticity terms, 1% increase in the HH income leads to 0.185% increase in the VMT at the HH level. This result agrees with the magnitude effect of WKRTOSIZ variable, i.e. the ratio of number of workers in the HH over the whole HH size. The larger this ratio, the higher the travel-related activities (elasticity of 0.170%), suggesting that income plays an important role in the travel behavior of the HH. Finally, HHs which consist of one adult only are found to drive less compared to HHs with other life cycle classifications (0.209%), a result which is reasonable in the context that the higher the number of people in a HH, the higher the number of trips generated.

4.5. Asymptotic t-test of equality of individual coefficients between the national and the local model

The asymptotic t-test of equality of individual coefficients between the national and the local model is conducted in order to identify the similarities between the models, in statistical terms. The author is interested in identifying whether the combination of variables for each model (fuel efficiency model, and VMT model) is the same between the two regions (Nation vs. Iowa). In second place, the power of this test is to identify whether there are statistically significant differences in the estimated coefficients for the common variables between the two regions.

The estimation results of the asymptotic t-test of equality of individual coefficients between the national and the local model are presented in Table 4-7. Also, Table 4-7 serves as a useful basis for comparing the combination of variables for the two models (fuel efficiency model and VMT model) between the two regions (Nation vs. Iowa). The asymptotic t-test of equality of individual coefficients is conducted solely for those variables which are common in both specifications (otherwise noted as N/A).

Table 4-7: Asymptotic *t*-test of equality of individual coefficients

	Model		Coefficient		Variance		Asymptotic	Statistically
Variable	**Nation**	**Iowa**	**Nation**	**Iowa**	**Nation**	**Iowa**	**t-test**	**Significant?**
Average vehicle fuel efficiency at the HH level model								
LARGE	✓	✓	-0.182	-0.103	1.46E-06	0.00454	1.346	✓
LVMT	✓	✓	0.103	0.072	6.68E-06	5.09E-05	-9.389	✓
VEHOWNM2	✓	✓	-0.0002	-0.0004	2.4E-10	6.08E-09	5.030	✓
HYBRID	✓	✓	0.014	-0.035	5.12E-06	0.000107	3.326	✓
WRKTOSIZ	✓	✓	0.036	0.034	2.72E-06	6.66E-05	-4.048	✓
VEHAGE	✓	✓	-0.004	-0.004	1.66E-08	4.97E-07	5.578	✓
LMSASIZ	-	✓	-	0.002	-	2.23E-07	N/A	N/A
NOSUB	✓	-	0.0105	-	5.65E-06	-	N/A	N/A
LAND	✓	-	-0.009	-	2.08E-06	-	N/A	N/A

Variable	Model		Coefficient		Variance		Asymptotic t-test	Statistically Significant?
	Nation	**Iowa**	**Nation**	**Iowa**	**Nation**	**Iowa**		
URB	✓	-	0.018	-	1.49E-06	-	N/A	N/A
RAIL	✓	-	0.012	-	1.33E-06	-	N/A	N/A
TAX	✓	-	0.036	-	2.06E-05	-	N/A	N/A
HOMEOWN	✓	-	-0.013	-	9.79E-07	-	N/A	N/A
ONEADULT	✓	-	0.012	-	1.81E-06	-	N/A	N/A
Vehicle-miles traveled at the HH level model								
LEFFC	✓	✓	2.291	2.759	0.008019	0.137095	-4.952	✓
NOSUB	✓	✓	-0.453	-0.509	4.35E-05	0.001487	12.556	✓
ONEADULT	-	✓	-	-0.202	-	0.001113	N/A	N/A
URB	✓	✓	-0.141	-0.148	2.55E-05	0.000541	6.077	✓
LINCOME	✓	✓	0.14	0.178	1.07E-05	0.000445	-8.201	✓
LMSASIZ	✓	✓	-0.004	-0.009	1.12E-07	4.57E-06	4.155	✓
LARGE	✓	✓	0.552	0.381	0.000228	0.001552	-8.477	✓
WRKTOSIZ	✓	✓	0.134	0.165	6.7E-05	0.001474	-4.069	✓
VEHOWNM2	✓	-	-0.002	-	6.51E-09	-	N/A	N/A
TAX	✓	-	-0.187	-	0.000351	-	N/A	N/A
LAND	✓	-	0.02	-	3.62E-05	-	N/A	N/A
RAIL	✓	-	-0.051	-	2.48E-05	-	N/A	N/A
DRVRCNT2	✓	-	0.218	-	1.66E-05	-	N/A	N/A

The results presented in Table 4-7 yield two different sets of conclusions. First, it is evident that each model has a similar set of combinations between the national and the regional model. Particularly for the fuel efficiency model, the two specifications share 6 common variables, whereas the national model has the NOSUB, ONEADULT and HOMEOWN variables, which were not found to be statistically significant at the local level, and also some location-specific variables (LAND, URB, RAIL, and TAX) which displayed no variation in the Iowa dataset. On the other hand, the fuel efficiency model for Iowa provides additional information through the variable LMSASIZ which is not present in the national model. Similarly, for the VMT model, the two specifications share 7 common variables, whereas the national model has the VEHOWNM2 and DRVRCNT2 variables, which were not found to be statistically significant at the local level, and also some location-specific variables (TAX, LAND, and RAIL) which displayed no variation in the Iowa dataset.

Regarding the actual estimation results of the asymptotic t-test, all differences are statistically significant at 99.90%, apart from the LARGE variable whose difference was found to be statistically significant at 90%.

However, it should be noted that the datasets at each level of analysis are not independent, thus there is an additional term that needs to be accounted for in the asymptotic t-ratio estimation. This term represents the covariance between the variable's estimated coefficients at the national and the local level. In this analysis, the estimated asymptotic t-ratio does not account for this term, thus the computed t-ratios presented in Table 4-7 are slightly different. However, given that the computed t-ratios are quite large, it is assumed that, even if the covariance is accounted for in the t-ratio estimation, the differences in the variables coefficients between the national and the local model would still be statistically significant. The only variable whose difference may not be statistically significant is the LARGE variable, whose value is very close to the 1.282 cut-off value at 0.10 level of significance.

4.6. Summary

This chapter identified the factors that mostly affect average vehicle fuel efficiency (as a proxy of fuel consumption, which is representative of the fuel tax pricing option) and VMT (representative of the VMT fee pricing option). The author also examined whether the inferences from a national model can be adopted at the local level (for example, Iowa).

4.6.1. Fuel tax model

The results that the two national models show that the VMT model has a higher R^2 value, suggesting a better fit of the model to the 2009 NHTS observations, than the fuel efficiency model.

The model specification also shows that a variety of factors affect travel behavior, represented by the purchase of high fuel efficiency vehicles, or the actual vehicle-miles traveled at the household level of analysis. This set of factors includes vehicle-specific characteristics, geographic characteristics, and socioeconomic characteristics.

By linking the average vehicle fuel efficiency to the average vehicle fuel consumption, the fuel efficiency model suggests that HHs that own hybrid vehicles are less likely to be affected by the current pricing option of fuel tax. Interestingly enough, HHs which own only one vehicle, and do not have the alternative of a second vehicle, are more likely to drive a vehicle of higher fuel efficiency than those HHs which own more than one. The vehicle ownership period has a negative effect on the fuel efficiency under which the vehicle is performing, suggesting that HHs which own their vehicles for a long time are most likely to drive less fuel efficient vehicles. The information provided by this variable is related to the HH income and the vehicle age. Regarding the vehicle type, HHs which own large vehicles, namely SUVs and pick-up trucks – are expected to consume more fuel, thus have higher fuel-related expenditures. Moreover, the endogenous variable of VMT suggests that HHs that generate a large amount of trips are more likely to invest on a fuel efficient vehicle, in order to balance the fuel expenditures which results from their high travel activity.

The location of the HH also plays a very important role on what level of fuel efficiency vehicles they operate. More precisely, HHs located in residential areas operate vehicles of lower fuel efficiency (thus consume more per mile traveled) than their counterparts, located in non-residential areas. The same applies to HHs located in rural areas, or in areas where there is no availability of an alternative mode of travel. Additionally, the effect of the actual taxation in effect also affects the choice of fuel efficient vehicles; it was found that HHs located in areas where the total of the federal and state tax is higher tend to operate vehicles of higher fuel efficiency, leading to reduced fuel-related expenditures per mile traveled.

In terms of socio-economic characteristics, it is inferred that the higher the workers to household size ratio, the more likely is the household to own and operate a fuel efficient vehicle. Additionally, HHs which own the housing unit are less likely to own a fuel efficient vehicle thus are expected to carry a greater portion of the fuel tax burden than their counterparts who rent it. Finally, HHs which consist of only one adult, tend to drive vehicles of higher fuel efficiency than these which consist of 2+ adults.

4.6.2. VMT fee model

Regarding the VMT fee model specification, the statistically significant factors are also categorized in vehicle-, geography- and socio-economic-specific groups. According to the parameter estimations of the VMT model, HHs which own high fuel efficiency vehicles have higher VMT). The same applies to HHs which own large vehicles, i.e. pick-up trucks and SUVs. On the other hand, HHs that own their vehicles for a longer period tend to drive less compared to HHs with newly purchased vehicles. Also, HHs which own more than one vehicle are found to drive more, probably due to high travel activity that may be associated with increased number of drivers in the HH or increased number of HH members.

Location-wise, HHs located in rural areas drive less than those located in urbanized areas. On the other hand, the size of the MSA has a positive effect on the vehicle miles traveled. Due to the same reason of decentralized activities, it was found that HHs located in residential areas drive more. The current level of taxation has a negative effect on VMT, as expected, as HHs located in areas where taxation is higher choose to reduce their VMT in an attempt to reduce fuel consumption. Finally, the presence of rail in an area reduces the VMT of the surrounding HHs, since it serves as mode alternative and part of the generated trips are assigned to this alternative.

In terms of socio-economic attributes, HHs with higher number of drivers generate more trips, thus have a higher VMT value. This may also be related to number of workers, suggesting that HHs with higher WRKTOSIZ ratio drive more. In the same context, HHs of higher income are found to drive more because they have smaller cost constraints than their counterparts of low or medium income.

4.6.3. National vs. Local model

Regarding the second part of the analysis, it is of great importance to comment on two aspects of this comparison:

a. The contribution of the statistical methodology applied in the current analysis
b. The differences in the specification of the local models vs. the national ones.

The application of the 3SLS estimator is of utmost importance in order to address the significant issue of endogeneity. The application of simple multiple regression models would not yield realistic results, and also would lead to worse model fit.

Via the analysis of variance, it was found that the level of spatial analysis does not have a statistically significant effect on the average vehicle fuel efficiency at the household level, but it does have a statistically significant effect on the vehicle-miles traveled.

The development of a local model showed that the two levels of analysis share quite a few common variables in the model specification. The location-specific variables, as those were derived from the 2009 NHTS, did not participate in the local model specification due to zero variability of those variables in the Iowa data. Additionally, for the set of variables which participate in both the national and the local version of the same model, the asymptotic t-test of equality of individual coefficients showed that there was a statistically significant difference detected in the coefficient estimates of the all participating variables, suggesting that the effect magnitude of each variable on the average vehicle fuel efficiency and on the HH's VMT differs between the national and the local model.

Chapter 5: Conclusions, Limitations and Recommendations

5.1. Introduction

The objective of the current analysis is two-fold. First, it is of interest to identify which geographic, demographic and socioeconomic groups are most likely to be affected under each pricing option, by developing statistical models for each alternative. The second research objective is to identify whether there are detectable and statistically significant differences between a nationally developed model and a local one for Iowa. Summarizing and extending the previous discussion, the implications of the findings of this thesis are presented in the Sections 5.2.1 and 5.2.2.

5.2. Implication of findings

The two following sections discuss the main conclusions that can be drawn from the model findings, as those were presented in Chapter 4: Methodology and Results. This discussion is conducted within the context of each pricing alternative, in a bid to interpret the model findings into policy recommendations.

5.2.1. Fuel Tax Pricing Option vs. VMT fee Pricing Option

The results of the average vehicle fuel efficiency model suggest that there are particular social sub-groups that operate vehicles of lower fuel efficiency at the HH level. It is very interesting that some of these groups also tend to have higher VMT at the HH level, thus it may be assumed that these groups will be negatively affected under both pricing options (however, the magnitude effect is defined by the actual coefficients). Therefore, at the national level, these common sub-groups include HHs which have the following characteristics:

- Operate large vehicles, namely pickup trucks or SUVs
- Own more than one vehicle
- Own vehicles for a longer time period
- Are located in rural areas
- Are located in an MSA with no heavy rail availability in the area,
- Are located in states with lower fuel taxation (sum of state plus federal tax)
- Are located in residential areas versus mixed land use areas

On the other hand, there were some particular sub-groups which were found to operate vehicles of lower fuel efficiency at the HH level, and there was no statistically valid information regarding their VMT. These groups are expected to shoulder greater of the fuel tax burden, as their relation to the average vehicle fuel efficiency at the HH level via the model specification is decreasing. This means that currently, HHs which meet one of the above criteria tend to operate vehicles of lower fuel efficiency, thus consume more fuel for the same distance traveled, and therefore have higher fuel-related expenditures. Most of the aforementioned sub-groups may be indirectly related to particular HH income levels. Although it would not be appropriate to over-state the model results, it may be generally assumed that these sub-groups are of limited purchasing power. These groups consist of HHs which:

- Own and operate conventional vehicles, versus hybrid vehicles or vehicles of alternative fuel technology
- Operate older vehicles
- Drive less at the HH level (lower VMT)

- Have low workers to HH size ratio
- Consist of two or more adults
- Own the housing unit they reside in

Similarly, there are some groups which are found to have higher VMT at the HH level, although the fuel efficiency model did not yield any statistically valid results regarding their average vehicle fuel efficiency at the HH level. The HHs which belong in one or more of the aforementioned sub-groups generate more trips at an annual basis, thus have a higher VMT at the HH level. Each sub-group mentioned in the previous list displays, due to prevailing circumstances, a higher need to generate more trips, or to be more accurate, to travel more miles. Therefore, in the event of implementing a new VMT fee policy, it is more likely that these subgroups will either carry a greater partition of the financial burden associated to the particular policy, or will be induced to reduce their travel in terms of VMT. Income-wise, there is no particular pattern across the sub-groups, however each distinct group may be disproportionally affected under this pricing alternative, compared to their counterparts that bear the opposite characteristics.

These groups consist of HHs which:

- Own vehicles of higher fuel efficiency
- Are located in smaller MSAs
- Have a higher number of licensed drivers
- Have a higher average income
- Have a higher workers to HH size ratio

The objective of this thesis is not to provide support for either of the two pricing alternatives; on the contrary, it is to provide statistically supported evidence of what social groups will most likely have increased fuel expenditures in each case. The information provided by the models of this thesis is highly valuable to policymakers in order to identify whether these alternatives are equitable among the various social subgroups. It is evident that some subgroups are affected more than others, and to a different extent, thus the conclusions presented in this thesis may assist policymakers in designing and implementing a pricing option that does not have strong negative impacts on particularly vulnerable social groups. However, the herein analysis requires further steps in order to interpret these results in a price-wise context. As it is further discussed in Sections 5.3 and 5.4, some enhancements may improve the findings of this analysis, which currently do not provide either qualitative or quantitative measures of the impact of a particular fee level to the travel behavior of various social groups.

5.2.2. National vs. Local model

Regarding the second part of the analysis, it is of great importance to comment on two aspects of this comparison:

- The validity of the endogenous-free model specification via the 3SLS estimator, and
- The differences in the specification of the local models vs. the national ones.

In terms of the first aspect, the application of the 3SLS estimator is of high importance in order to address the significant issue of endogeneity between the average vehicle fuel efficiency and the vehicle miles traveled at the HH level. This estimator yields endogeneity-free coefficients which capture the interdependencies between the two dependent variables more successfully.

Regarding the second aspect, the differences between the national and local model lie in three different levels, suggesting that, despite the similarities, the development of a distinct local model is statistically supported.

- There are statistically significant differences in the VMT at the HH level, suggesting that VMT in Iowa exhibit different trend than the national average VMT.
- The two levels of analysis share quite a few common variables in the model specification, but the location-specific variables did not participate in the local model development.
- Based on the asymptotic t-test of equality of individual coefficients, there is a statistically significant difference detected in the coefficient estimates of all variables, suggesting that the effect magnitude of each variable on the average vehicle fuel efficiency at HH level and on the HH's VMT differs between the national and the local model.

5.3. Study limitations

The main limitation of this study pertains to the results under the VMT fee pricing option. The results of the VMT model specification are not totally able to capture the change in travel behavior that the implementation of a VMT fee may cause. As a new policy measure, the effect of the implementation of the VMT fee on the behavior of travelers does not only depend on the actual fee. It also depends significantly on the exact implementation process, such as the fee collection mechanism. For example, travel behavior may be affected in a different way depending on when and how the fees are collected. A monthly invoice will most probably have different effect on VMT than a bi-weekly pattern, which mostly resembles the current fuel purchase at-the-pump.

With respect to the model specification, state is the highest level of analysis that this survey allows for, thus there is no detailed information for variables which may vary within the state boundaries. For example, the availability of transit systems, other than the heavy rail, would be a useful piece of information to include in both models. However such information is expected to vary significantly within the state but would possibly have greater explanatory power in the local model specification.

Furthermore, the analysis focused solely on 2009, due to data availability. This prevented the author from identifying time-specific factors that affect travel behavior, as this is captured via average vehicle fuel efficiency and VMT at the HH level, and also sets some limitations in the models transferability over time.

Finally, significant changes in future transportation conditions, such as higher future market penetration of alternative fuel or hybrid vehicles, or new modes such as high-speed rail may modify the findings and the models might need modifications to capture these future conditions.

5.4. Recommendations for future research

While this thesis provided insights on the equity impacts of the two policy mechanisms, a few recommendations for future research are provided below.

In order to obtain a more realistic perspective of how a VMT fee would affect daily travel, future research should focus on designing and conducting an experimental analysis.

This has been done in the past but not extensively. Such an experiment should simulate the suggested VMT fee implementation mechanism, which may include an on-board unit for tracking VMT, and an invoice sent at regular times to the experimental subjects, or a similar scheme. Such an experiment, if designed properly, taking into account most of the parameters, would better simulate changes in travel behavior due to the fee implementation.

Also, it is recommended that future work combines the aggregate NHTS data information with more disaggregate data (for example at the local level) in order to capture the variance of localized variables, and make inferences regarding the local characteristics of a state.

Moreover, it is suggested that future research analyzes the same research question using time-series data, in order to capture the time effect on travel behavior. This will be particularly interest in identifying the long-term equity impacts of each alternative, which is more significant in decision-making. It is also suggested that the same type of analysis is performed for other states as well, in order to validate the conclusions reached in this thesis regarding the national and local model specification.

Finally, it is recommended that future research studied these research questions at the person-level, to allow for distinct travel behaviors with one HH. The HH concept may be considered outdated today, where people who are not related to each other co-reside in the same housing unit, but display distinct travel behavior, and are bear individually their share of the HH’s fuel expenditures

.

Bibliography

Agrawal, A. W., Dill, J., & Nixon, H. (2010). Green Transportation taxes and fees: a survey of public preferences in California. *Transportation Research Part D, 15*, 189-196.

American Petroleum Institute. (n.d.). Diesel Taxes: Combined Local, State, and Federal (cents per gallon) - January 2012. Retrieved from http://www.api.org/Oil-and-Natural-Gas-Overview/Industry-Economics/~/media/A375B82CC4184656A093C6168A1DD08E.ashx

American Petroleum Institute. (n.d.). Gasoline Taxes: Combined Local, State, and Federal (cents per gallon) - January 2012. Retrieved from http://www.api.org/Oil-and-Natural-Gas-Overview/Industry-Economics/~/media/21EBD0B62EBA42B1965EE82EFFB6585D.ashx

American Petroleum Institute. (n.d.). Notes to state motor fuel excise and other taxes. Retrieved from http://www.api.org/Oil-and-Natural-Gas-Overview/Industry-Economics/~/media/Files/Statistics/State_Motor_Fuel_Excise_Tax_Update.ashx

Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand.* MIT Press.

Bento, A. M., Goulder, L. H., Jacobsen, M. R., & von Haefen, R. H. (2009). Distributional and Efficiency impacts of increased US gasoline taxes. *American Economic Review, 99*(3), 667-699.

Boos, M. A., & Moruza, A. K. (2008). *Vehicle Miles Traveled (VMT) Tax: An Alternative to the Gas Tax for Generating Highway Revenue.* Research Synthesis Bibliography No. 19, Virginia Department of Transportation.

Busse, M. R., Knittel, C. R., & Zettelmeyer, F. (2009). *Pain at the pump: the differential effect of gasoline prices on new and used automobile markets.* NBER Working Paper 15590, National Bureau of Economic Research.

Cobb, G. W. (1998). *Introduction to design and analysis of experiments.* Springer-Verlag New York, Inc., and Key College Publishing.

Council of Academies of Engineering and Technological Sciences. (1995). *The Role of Technology in 6 Environmentally Sustainable Development.* Washington D.C.: The National Academies Press.

Deakin, E., & Harvey, G. (1996). *Transportation Pricing Strategies for California: an Assessment of Congestion, Emissions, Energy, and Equity Impacts.* University of California Transportation Center .

Forkenbrock, D. J., & Kuhl, J. G. (2002). *A new approach to assessing road user charges.* Public Policy Center, University of Iowa.

Fullerton, D., & West, S. E. (2002). Can taxes on cars and on gasoline mimic an unavailable tax on emissions? *Journal of Environmental Economics and Management, 43*, 135-157.

Greene, D. L. (2011). What is greener than a VMT tax? The case of an indexed energy user fee to finance US surface transportation. *Transportation Research Part D, 16*, 451-458.

Greene, D. L., Patterson, P. D., Singh, M., & Li, J. (2005). Feebates, rebates and gas-guzzler taxes: a study of incentives for increased fuel economy. *Energy Policy, 33*, 757-775.

Hanley, P. F., & Kuhl, J. G. (2011). National Evaluation of Mileage-Based Charges for Drivers: Initial Results. *Transportation Research Record: Journal of the Transportation Reseach Board, No. 2221*, 10-18.

International Atomic Energy Agency. (2005). *Energy indicators for sustainable development: guidelines and methodologies.* Vienna.

Kayser, H. A. (2000). Gasoline demand and car choice: estimating gasoline demand using household information. *Energy Economics , 22*, 331-348.

Krupnick, A., Harrington, W., & Alberini, A. (2001). Public support for pollution fee policies for motor vehicles with revenue recycling: survey results. *Regional Science and Urban Economics, 31*, 505-522.

Larsen, L., Burris, M., Pearson, D., & Ellis, P. (2012). Equity evaluation of Vehicle Miles Traveled Fees in Texas. *TRB 2012 Annual Meeting CD-ROM.*

Li, S., Linn, J., & Muehlegger, E. (2011). *Gasoline taxes and consumer behavior.* NBER Working Paper 17891, National Bureau of Economic Research.

Lindsey, M., Schofer, J. L., Durango-Cohen, P., & Gray, K. A. (2011). The effect of residential location on vehicle miles of travel, energy consumption and greenhouse gas emissions: Chicago case study. *Transportation Research Part D, 16*, 1-9.

Litman, T. (1999). Distance-based charges; a practical strategy for more optimal vehicle pricing. Retrieved from www.vtpi.org

Litman, T., Komanoff, C., & Howell, Esq., D. (1998). *Road Relief: Tax and Pricing Shifts for a Fairer, Cleaner, and Less Congested Transportation System in Washington State.* Energy Outreach Center.

McMullen, B. S., Zhang, L., & Nakahara, K. (2010). Distributional impacts of changing from a gasoline tax to a vehicle-mile tax for light vehicles: a case study for Oregon. *Transport Policy*, *17*, 359-366.

Moghadam, A. K. (2011). Abatement cost functions for incentive-based policies in the automobile market. *Transportation Research Part D, 16*, 579-585.

National Surface Transportation Infrastructure Financing Commission. (2009). *Paying our way: a new framework for transportation finance.*

National Surface Transportation Policy and Revenue Study Commission. (2007). *Report of the National Surface Transportation Policy and Revenue Study Commission: Transportation for Tomorrow*.

NHTSA. (n.d.). *CAFE - Fuel Economy*. Retrieved from http://www.nhtsa.gov/fuel-economy

NHTSA. (n.d.). *Historical Passenger Car Fleet Average Characteristics*. Retrieved from http://www.nhtsa.gov/cars/rules/CAFE/HistoricalCarFleet.htm

Obama announces 54.5 mpg CAFE standard by 2025. (n.d.). Retrieved from Popular Mechanics: http://www.popularmechanics.com/cars/news/fuel-economy/obama-announces-54-6-mpg-cafe-standard-by-2025

Peters, J. R., & Gordon, C. (2009). *Analysis of Alternative Funding Sources.*

Poterba, J. M. (1990). *Is the gasoline tax regressive?* NBER Working Paper No. 3578, National Bureau of Economic Research.

Pozdena, R. J. (1995). *Where the rubber meets the road: Reforming California's Roadway System.* Policy Study No. 191, Reason Foundation.

Rizzo, L., Swain, S., Freedman, M., Hubbell, K., Kang, J., Morganstein, D., et al. (2009). Tak N: Weighting Report. Retrieved from http://nhts.ornl.gov/2009/pub/WeightingReport.pdf

Rufolo, A. M. (2011). Cost Estimates for Collecting VMT fees. *TRB 2011 Annual Meeting.*

Schank, J., & Rudnick-Thorpe, N. (2011). End of the Highway Trust Fund? Long-Term Options for Funding Federal Surface Transportation. *Transportation Research Record: Journal of the Transportation Research Board, No. 2221*, 1-9.

Sisiopiku, V. P., & Waid, J. C. (2007). The potential of the vehicle mileage road user fee as an alternate financing option for state highways. *TRB 2007 Annual Meeting CD-ROM.*

Sisiopiku, V. P., Waid, J. C., Rizk, T., McLeod, R., & Robins, W. (2006). *Alternate Financing Sources for Alabama Highways.* UTCA Report Number 05114, University Transportation Center for Alabama.

Southeast Michigan Council of Governments. (2006). Transportation Funding: A critical economic issue for Southeast Michigan. *Proceedings of the Transportation Funding Discussion Series.*

The World Conservation Union. (2006). *The future of sustainability: Re-thinking environment and development in the Twenty-first Century.* Report of the IUCN Renowned Thinkers Meeting, 29-31 January 2006.

Tollng and Pricing Program: Value Pricing Pilot Program. (n.d.). Retrieved 2012, from US Department of Transportation, Federal Highway Administration: http://www.ops.fhwa.dot.gov/tolling_pricing/value_pricing/index.htm

TRB Committee for the Study of the Long-Term Viability of Fuel Taxes for Transportation Finance. (2006). *Special Report 285: The Fuel Tax and Alternatives for Transportation Funding.* Transportation Research Board of the National Academies.

TRB Committee on Equity Implications of Evolving Transportation Finance Mechanisms. (2011). *Special Report 303: Equity of Evolving Transportation Finance Mechanisms.*

U.S. Department of Transportation, F. H. (n.d.). 2009 National Household Survey: Codebook. Retrieved from http://nhts.ornl.gov/2009/pub/Codebook.pdf

U.S. Department of Transportation, F. H. (n.d.). 2009 National Household Travel Survey: User's Guide (Version 2). Retrieved from http://nhts.ornl.gov/2009/pub/UsersGuideV2.pdf

U.S. Department of Transportation, Federal Highway Administration. (n.d.). 2009 National Household Travel Survey. Retrieved from www.nhts.ornl.gov

U.S. Department of Transportation, Federal Highway Administration. (n.d.). *Introduction to the 2009 NHTS*. Retrieved from NHTS: National Household Travel Survey, Our Nation's Travel: http://nhts.ornl.gov/introduction.shtml

U.S. Department of Transportation, NHTSA. (2011). *Summary of Fuel Economy Performance.*

U.S. Environmental Protection Agency. (n.d.). Sustainability: Basic Information. Retrieved from http://www.epa.gov/sustainability/basicinfo.htm

United Nations. (1987). *G. H. Brundtland Report of the World Commission on Environment and Development: Our Common Future*. New York.

Vacation 2 USA. (n.d.). Retrieved 2011, from Vacation 2 USA: http://www.vacation2usa.com/i2map_iowa.png

Vera, I. A., Langlois, L. M., Rogner, H. H., Jalal, A. I., & Toth, F. L. (2005). Indicators for sustainable energy development: an initiative by the International Atomic Energy Agency. *Natural Resources Forum, 29*, 274-283.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2011). *Statistical and econometric methods for transportation data analysis* (2 ed.). Taylor and Francis Group, LLC.

Weatherford, B. A. (2011). Distributional Implications of Replacing the Federal Fuel Tax with Per Mile User Charges. *Transportation Research Record: Journal of the Transportation Research Boar, No. 2221*, 19-26.

West, S. E. (2004). Distributional effects of alternative vehicle pollution control policies. *Journal of Public Economics*, *88*, 735-757.

Whang, J., Jinga, Y. Y., Zhanga, C. F., & Zhaoa, J. H. (2009). Review on multi-criteria decision analysis aid in sustainable energy decision-making. *Renewable and Sustainable Energy Reviews, 13*, 2263-2278.

Whitty, J. M., & Imholt, B. (2005). *Oregon's Mileage Fee Concept and Road User Fee Pilot Program: Report to the 73rd Oregon Legislative Assembly*.

Whitty, J., Svadlenak, J., & Capps, D. (2006). *Public Involvement and Road User Charge Development: Oregon's Experience.* Oregon Department of Transportation.

Zhang, L., & Lu, Y. (2012). Marginal-cost vehicle mileage fee. *TRB 2012 Annual Meeting CD-ROM.*

Zhang, L., McMullen, B. S., Valluri, D., & Nakahara, K. (2009). The short- and long-run impacts of vehicle mileage fee on income and spatial equity. Retrieved from http://wiki.umd.edu/lei/images/f/fe/Zhang2009b.pdf

Appendix 1A: Descriptive Statistics – Nation

Variable mnemonic	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
ALT	HH owns alternative fuel vehicle	1.04E-03	3.22E-02	0	1	142942	0
AUTO	Vehicle type HH owns: automobile/car/station wagon	0.723211	0.447413	0	1	142942	0
BESTMILE	Best estimate of annual miles	20608.2	16661.1	0	409006	138056	4886
CENSUS_R	Census region classification for home address	2.65741	1.00233	1	4	142942	0
COMM	HH owns commercial license plate vehicle	4.38E-02	0.204692	0	1	142942	0
CWEIADMP	EIA derived miles per gasoline-equivalent gallon estimate (weighted average) – outliers removed	21.0234	4.1055	10.8916	32.1742	122703	20239
DIESEL	HH owns diesel-fueled vehicle	2.38E-02	0.152275	0	1	142942	0
DRVRCNT	Number of drivers in HH	1.8388	0.781971	0	9	142942	0
DRVRCNT2	Number of drivers in HH - outliers removed	1.8229	0.746648	0	4	142395	547
ELECTRIC	HH owns electric vehicle	7.22E-05	8.50E-03	0	1	142942	0
FED	Federal fuel tax	0.184	1.36E-15	0.184	0.184	142942	0
GSCOST	Fuel cost in nominal US dollars per gasoline equivalent gallon	3.04004	0.130693	1.57	4.53	141522	1420
HBPPOPDN	Population per sq. mile - Block group	4507.14	6087.01	50	30000	142941	1
HBRESDN	Housing units per sq. mile - Block	2077.83	3692.33	50	30000	142941	1

Variable mnemonic	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
	group						
HH_HISP	Hispanic status of HH respondent	0.107113	0.309258	0	1	142382	560
HH_RACE	Race of HH respondent	2.98492	11.9093	1	97	141803	1139
HHFAMINC	Derived total HH income	10.747	5.5464	1	18	131871	11071
HHSIZE	Count of HH members	2.53319	1.3965	1	14	142942	0
HHSIZE2	Count of HH members - outliers removed	2.4646	1.27984	1	6	141550	1392
HHSTFIPS	State FIPS for HH address	27.7307	15.9778	1	56	142942	0
HHVEHCNT	Count of HH vehicles	2.04097	1.10895	1	27	142942	0
HHVHCNT2	Count of HH vehicles - outliers removed	1.99946	1.01146	1	6	140925	2017
HIGH	Total HH income: $40,001-$60,000	0.1856	0.388785	0	1	131871	11071
HOMEOWN	Housing unit owned	0.71289	0.452415	0	1	142942	0
HOMETYPE	Type of housing unit	1.70102	3.64956	1	97	142661	281
HOUSEID	HH eight-digit ID number	4.49E+07	1.44E+07	2.00E+07	7.00E+07	142942	0
HTEEMPDN	Workers per square mile living in Tract	1166.32	1459.21	25	5000	142941	1
HTPPOPDN	Population per sq. mile - Tract level	3962.8	5662.11	50	30000	142941	1
HYBRID	HH owns hybrid vehicle	4.31E-02	0.203052	0	1	142942	0
INCOME	Derived total HH income - continuous	54142.7	31589.8	2500	100000	131871	11071
INCPCAP	Income per capita	26245.8	19020.6	416.667	100000	130518	12424
LAND	Land use: residential	0.873524	0.332386	0	1	142942	0
LARGE	HH owns large vehicle (i.e. sports utility vehicle or pickup truck)	0.51471	0.499785	0	1	142942	0
LCONSUM	Vehicle Fuel Consumption - Log	6.64424	0.956228	0	11.1197	141653	1289

Variable mnemonic	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
LEFF	EIA derived miles per gasoline-equivalent gallon estimate (weighted average) - Log	3.04202	0.231327	1.8563	4.76217	137902	5040
LEFFC	EIA derived miles per gasoline-equivalent gallon estimate (weighted average) – Log – outliers removed	3.02616	0.199201	2.38799	3.47117	122703	20239
LEXP	Annual fuel expenditures in nominal US dollars (sum) - Log	7.75417	0.962492	0	12.3785	141653	1289
LGSCOST	Fuel cost in nominal US dollars per gasoline equivalent gallon	1.11097	4.23E-02	0.451076	1.51072	141522	1420
LIF_CYC	Life Cycle classification for the HH	5.09667	3.34456	1	10	142942	0
LINCOME	Derived total HH income - continuous - Log	10.65	0.811549	7.82405	11.5129	131871	11071
LMSASIZ	MSA population size for the HH home address - continuous - Log	11.247	5.77538	0	14.9141	131870	11072
LOW	Total HH income: $2,500-$20,000	0.174511	0.37955	0	1	131871	11071
LVMT	Best estimate of annual miles - Log	9.57729	-1.00737	3.56137	12.9215	138056	4886
MED	Total HH income: $20,001-$40,000	0.234878	0.423924	0	1	131871	11071
MGAS	HH owns motor gasoline vehicle	0.956303	0.204421	0	1	142942	0
MIDWEST	Census region classification for home address: Midwest	0.232301	0.422301	0	1	142942	0
MINOR	Race of HH respondent: African American, Black, American Indian, Alaskan Native, Native Hawaiian, other Pacific, Hispanic/Mexican	0.154322	0.361258	0	1	142942	0

Variable mnemonic	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
MSACAT	MSA category for the HH home address	2.38081	1.08078	1	4	142941	1
MSASIZ	MSA population size for the HH home address - continuous	1.54E+06	1.25E+06	0	3.00E+06	131870	11072
MSASIZE	MSA population size for the HH home address	4.26598	1.4685	1	6	142941	1
N_AUTO	Number of automobile/car/station wagon vehicles in HH	1.02249	0.806942	0	22	142942	0
N_AUTO2	Number of automobile/car/station wagon vehicles in HH - outliers removed	0.986876	0.74433	0	3	139121	3821
N_DIESEL	Number of diesel-fueled vehicles in HH	2.56E-02	0.173056	0	5	142942	0
N_ELECTR	Number of electric vehicles in HH	7.22E-05	8.50E-03	0	1	142942	0
N_HYBRID	Number of hybrid vehicles in HH	7.51E-02	0.289062	0	4	97276	45666
N_MGAS	Number of motor gasoline vehicles in HH	2.01943	1.09422	0	26	142942	0
N_MGAS2	Number of motor gasoline vehicles in HH - outliers removed	1.93183	0.950419	0	6	137726	5216
N_NGAS	Number of natural-gas fueled vehicles in HH	9.77E-04	3.15E-02	0	2	142942	0
N_OTHER	Number of other vehicles in HH	1.72077	1.46343	0	27	142942	0
N_OTHER2	Number of other vehicles in HH - outliers removed	1.61668	1.33777	0	6	137726	5216
N_PICKUP	Number of pickup truck vehicles in HH	0.365423	0.612907	0	9	142942	0

Variable mnemonic	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
N_PICUP2	Number of pickup truck vehicles in HH - outliers removed	0.33791	0.550025	0	2	137743	5199
N_SUV	Number of sports utility vehicles in HH	0.398131	0.601686	0	10	142942	0
N_SUV2	Number of sports utility vehicles in HH - outliers removed	0.381206	0.569743	0	2	138577	4365
N_VEHCO2	Number of commercial license plate vehicles in HH - outliers removed	5.28E-02	0.254275	0	2	138843	4099
N_VEHCOM	Number of commercial license plate vehicles in HH	6.26E-02	0.311334	0	11	141936	1006
NGAS	HH owns natural-gas fueled vehicle	9.69E-04	3.11E-02	0	1	142942	0
NORTH	Census region classification for home address: Northeast	0.166668	0.37268	0	1	142942	0
NOSUB	HH with only one vehicle	0.353514	0.478062	0	1	142942	0
NUMADLT	Count of adult HHMs at least 18 years old	1.92306	0.761604	1	10	142942	0
NUMADLT2	Count of adult HHMs at least 18 years old - outliers removed	1.89088	0.707135	1	4	140706	2236
ONEADULT	Life cycle classification of HH: one adult	0.286721	0.452232	0	1	142942	0
OTHER	Vehicle type HH owns: other	0.634985	0.481436	0	1	142942	0
PICKUP	Vehicle type HH owns: pickup truck vehicle	0.285873	0.451831	0	1	142942	0
RAIL	MSA has rail	0.263987	0.440794	0	1	142942	0
RATEXP	Fuel expenditure to HH income	9.24E-02	0.162679	0	10.208	130883	12059

Variable mnemonic	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
	ratio						
RATEXPC	Fuel expenditure to HH income ratio – corrected values	0.28523	3.98392	0	185.423	124064	18878
SGSTOTCS	Annual fuel expenditures in nominal US dollars (sum)	3349.42	3054.04	0	237630	141653	1289
SGSYRGAL	Annual fuel consumption in gasoline equivalent gallons (sum)	1102.65	992.11	0	67487	141653	1289
SOUTH	Census region classification for home address: South	0.377983	0.484885	0	1	142942	0
STATE	State/local fuel tax	0.311559	0.10973	8.00E-02	0.496	131871	11071
SUV	Vehicle type HH owns: sports utility vehicle	0.324484	0.468183	0	1	142942	0
TAX	Fuel tax (federal plus state/local)	0.495559	0.10973	0.264	0.68	131871	11071
TPM	Tax per mile cost	2.42E-02	7.66E-03	3.19E-03	0.105469	127480	15462
URB	Household in urban area	0.757218	0.428766	0	1	131870	11072
URBAN	Home address in urbanized area	1.83269	1.26006	1	4	142941	1
URBANSIZ	Size of urban area in which home address is located	4.31711	1.68917	1	6	142941	1
URBRUR	Household in urban/rural area	1.24108	0.427741	1	2	142941	1
VEHAGE	Age of vehicle in years (weighted average)	7.58189	4.55294	0	24	137902	5040
VEHOWNM2	How long vehicle(s) owned - Months (weighted average) - outliers removed	51.7591	36.874	0	186.95	130565	12377
VEHOWNMO	How long vehicle(s) owned - Months (weighted average)	53.8187	42.3274	0	623.211	137902	5040

Variable mnemonic	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
VHIGH	Total HH income: $60,001-$100,000	0.405012	0.490896	0	1	131871	11071
WEIADMPG	EIA derived miles per gasoline-equivalent gallon estimate (weighted average)	21.5297	5.3233	6.4	117	137902	5040
WEST	Census region classification for home address: West	0.223048	0.416292	0	1	142942	0
WGSTOTCS	Annual fuel expenditures in nominal US dollars (weighted average)	1776.89	1430.06	0	75692.8	137902	5040
WGSYRGAL	Annual fuel consumption in gasoline equivalent gallons (weighted average)	584.829	457.013	0	21264.6	137902	5040
WHITE	Race of HH respondent: white	0.789992	0.407315	0	1	142942	0
WRKCNT2	Number of workers in HH - outliers removed	1.11707	0.812443	0	3	140141	2801
WRKCOUNT	Number of workers in HH	1.15098	0.856864	0	6	142942	0
WRKTOSIZ	Number of workers to HH size ratio	0.502896	0.378755	0	1	140141	2801
WTHHFIN	Final HH weight	6143.02	8471.42	1.1709	53066.2	142942	0

Appendix 1B: Descriptive Statistics – Iowa

Variable	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
ALT	HH owns alternative fuel vehicle	2.85E-04	1.69E-02	0	1	3614	0
AUTO	Vehicle type HH owns: automobile/car/station wagon	0.731365	0.443311	0	1	3614	0
BESTMILE	Best estimate of annual miles	22336.4	16287.8	0	178644	3491	123
CENSUS_R	Census region classification for home address	2	0	2	2	3614	0
COMM	HH owns commercial license plate vehicle	3.33E-02	0.179454	0	1	3614	0
CWEIADMP	EIA derived miles per gasoline-equivalent gallon estimate (weighted average) – outliers removed	20.4663	3.73128	10.9	32.1739	3170	444
DIESEL	HH owns diesel-fueled vehicle	4.53E-02	0.208014	0	1	3614	0
DRVRCNT	Number of drivers in HH	1.83621	0.72099	0	6	3614	0
DRVRCNT2	Number of drivers in HH - outliers removed	1.82291	0.691325	0	4	3601	13
ELECTRIC	HH owns electric vehicle	0	0	0	0	3614	0
FED	Federal fuel tax	0.184	1.39E-16	0.184	0.184	3614	0
GSCOST	Fuel cost in nominal US dollars per gasoline equivalent gallon	2.9514	4.26E-02	2.315	3.49	3591	23
HBPPOPDN	Population per sq. mile - Block group	2087.89	2761.48	50	30000	3614	0
HBRESDN	Housing units per sq. mile - Block group	912.344	1179.96	50	7000	3614	0
HH_HISP	Hispanic status of HH respondent	3.03E-02	0.171301	0	1	3606	8
HH_RACE	Race of HH respondent	1.36726	5.227	1	97	3601	13
HHFAMINC	Derived total HH income	10.4109	4.92558	1	18	3367	247

Variable	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
HHSIZE	Count of HH members	2.40781	1.2786	1	9	3614	0
HHSIZE2	Count of HH members - outliers removed	2.36517	1.20795	1	6	3583	31
HHSTFIPS	State FIPS for HH address	19	3.55E-15	19	19	3614	0
HHVEHCNT	Count of HH vehicles	2.43828	1.33086	1	12	3614	0
HHVHCNT2	Count of HH vehicles - outliers removed	2.36437	1.17066	1	6	3554	60
HIGH	Total HH income: $40,001-$60,000	0.237705	0.42574	0	1	3367	247
HOMEOWN	Housing unit owned	0.758363	0.428134	0	1	3614	0
HOMETYPE	Type of housing unit	1.5378	3.68979	1	97	3607	7
HOUSEID	HH eight-digit ID number	4.36E+07	1.46E+07	2.00E+07	7.00E+07	3614	0
HTEEMPDN	Workers per square mile living in Tract	642.938	1033.33	25	5000	3614	0
HTPPOPDN	Population per sq mile - Tract level	1488.69	2190.58	50	17000	3614	0
HYBRID	HH owns hybrid vehicle	6.40E-02	0.244788	0	1	3614	0
INCOME	Derived total HH income - continuous	51441	27581.4	2500	100000	3367	247
INCPCAP	Income per capita	24664.9	15455.9	500	100000	3337	277
LAND	Land use: residential	0.831981	0.373935	0	1	3614	0
LARGE	HH owns large vehicle (i.e. sports utility vehicle or pickup truck)	0.676784	0.658483	0	2	3614	0
LCONSUM	Vehicle Fuel Consumption - Log	6.83177	0.925964	0	9.90872	3594	20
LEFF	EIA derived miles per gasoline-equivalent gallon estimate (weighted average) - Log	3.01216	0.213186	2.14007	4.06851	3488	126
LEFFC	EIA derived miles per gasoline-equivalent gallon estimate (weighted average) – Log – outliers removed	3.0021	0.1836	2.38876	3.47116	3170	444
LEXP	Annual fuel expenditures in nominal US dollars (sum) - Log	7.91066	0.946837	0	11.1533	3594	20
LGSCOST	Fuel cost in nominal US dollars per	1.08218	1.39E-02	0.83941	1.2499	3591	23

Variable	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
	gasoline equivalent gallon						
LIF_CYC	Life Cycle classification for the HH	5.20609	3.40869	1	10	3614	0
LINCOME	Derived total HH income - continuous - Log	10.649	0.722012	7.82405	11.5129	3367	247
LMSASIZ	MSA population size for the HH home address - continuous - Log	5.63116	6.11296	0	13.5278	3367	247
LOW	Total HH income: $2,500-$20,000	0.141642	0.348734	0	1	3367	247
LVMT	Best estimate of annual miles - Log	9.69461	1.01012	0	12.0931	3491	123
MED	Total HH income: $20,001-$40,000	0.271953	0.445031	0	1	3367	247
MGAS	HH owns motor gasoline vehicle	0.92993	0.255301	0	1	3614	0
MIDWEST	Census region classification for home address: Midwest	1	0	1	1	3614	0
MINOR	Race of HH respondent: African American, Black, American Indian, Alaskan Native, Native Hawaiian, other Pacific, Hispanic/Mexican	1.59E-02	0.125205	0	1	3614	0
MSACAT	MSA category for the HH home address	3.54755	0.497802	3	4	3614	0
MSASIZ	MSA population size for the HH home address - continuous	116572	174875	0	750000	3367	247
MSASIZE	MSA population size for the HH home address	3.95443	2.2892	1	6	3614	0
N_AUTO	Number of automobile/car/station wagon vehicles in HH	1.0759	0.828425	0	7	3614	0
N_AUTO2	Number of automobile/car/station wagon vehicles in HH - outliers removed	1.0207	0.741212	0	3	3507	107
N_DIESEL	Number of diesel-fueled vehicles in HH	4.78E-02	0.226275	0	3	3614	0

Variable	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
N_ELECTR	Number of electric vehicles in HH	0	0	0	0	3614	0
N_HYBRID	Number of hybrid vehicles in HH	0.111128	0.33869	0	2	2358	1256
N_MGAS	Number of motor gasoline vehicles in HH	2.3951	1.30579	0	12	3614	0
N_MGAS2	Number of motor gasoline vehicles in HH - outliers removed	2.20945	1.04551	0	6	3456	158
N_NGAS	Number of natural-gas fueled vehicles in HH	2.85E-04	1.69E-02	0	1	3614	0
N_OTHER	Number of other vehicles in HH	2.20679	1.62972	0	12	3614	0
N_OTHER2	Number of other vehicles in HH - outliers removed	2.00233	1.40877	0	6	3456	158
N_PICKUP	Number of pickup truck vehicles in HH	0.583997	0.795706	0	8	3614	0
N_PICUP2	Number of pickup truck vehicles in HH - outliers removed	0.484624	0.607223	0	2	3458	156
N_SUV	Number of sports utility vehicles in HH	0.344761	0.573505	0	5	3614	0
N_SUV2	Number of sports utility vehicles in HH - outliers removed	0.330731	0.548176	0	2	3497	117
N_VEHCO2	Number of commercial license plate vehicles in HH - outliers removed	4.10E-02	0.227986	0	2	3503	111
N_VEHCOM	Number of commercial license plate vehicles in HH	5.29E-02	0.323628	0	9	3589	25
NGAS	HH owns natural-gas fueled vehicle	2.85E-04	1.69E-02	0	1	3614	0
NORTH	Census region classification for home address: Northeast	0	0	0	0	3614	0
NOSUB	HH with only one vehicle	0.255191	0.436029	0	1	3614	0
NUMADLT	Count of adult HHMs at least 18 years old	1.82976	0.628277	1	6	3614	0
NUMADLT2	Count of adult HHMs at least 18 years old	1.80846	0.592829	1	4	3552	62

Variable	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
	- outliers removed						
ONEADULT	Life cycle classification of HH: one adult	0.281483	0.449785	0	1	3614	0
OTHER	Vehicle type HH owns: other	0.695973	0.460058	0	1	3614	0
PICKUP	Vehicle type HH owns: pickup truck vehicle	0.396531	0.489245	0	1	3614	0
RAIL	MSA has rail	0	0	0	0	3614	0
RATEXP	Fuel expenditure to HH income ratio	9.87E-02	0.183848	0	3.98829	3351	263
RATEXPC	Fuel expenditure to HH income ratio – corrected values	0.198555	2.33424	0	62.8342	3173	441
SGSTOTCS	Annual fuel expenditures in nominal US dollars (sum)	3746.97	3069.11	0	69795	3594	20
SGSYRGAL	Annual fuel consumption in gasoline equivalent gallons (sum)	1266.81	1006.13	0	20105	3594	20
SOUTH	Census region classification for home address: South	0	0	0	0	3614	0
STATE	State/local fuel tax	0.22	6.11E-16	0.22	0.22	3367	247
SUV	Vehicle type HH owns: sports utility vehicle	0.280253	0.449185	0	1	3614	0
TAX	Fuel tax (federal plus state/local)	0.404	5.55E-17	0.404	0.404	3367	247
TPM	Tax per mile cost	2.03E-02	4.36E-03	6.91E-03	4.75E-02	3253	361
URB	Household in urban area	0.579106	0.493775	0	1	3367	247
URBAN	Home address in urbanized area	2.48111	1.35417	1	4	3614	0
URBANSIZ	Size of urban area in which home address is located	4.32489	2.22697	1	6	3614	0
URBRUR	Household in urban/rural area	1.42417	0.494285	1	2	3614	0
VEHAGE	Age of vehicle in years (weighted average)	7.93991	4.16266	0	24	3488	126

Variable	Variable Description	Mean	Std. Dev.	Minimum	Maximum	Cases	Missing
VEHOWNM2	How long vehicle(s) owned - Months (weighted average) - outliers removed	51.7811	36.6158	0	186.843	3294	320
VEHOWNMO	How long vehicle(s) owned - Months (weighted average)	53.793	38.9039	0	384	3488	126
VHIGH	Total HH income: $60,001-$100,000	0.3487	0.476629	0	1	3367	247
WEIADMPG	EIA derived miles per gasoline-equivalent gallon estimate (weighted average)	20.8142	4.7967	8.5	58.4697	3488	126
WEST	Census region classification for home address: West	0	0	0	0	3614	0
WGSTOTCS	Annual fuel expenditures in nominal US dollars (weighted average)	1740.85	1501.85	4	62653.1	3488	126
WGSYRGAL	Annual fuel consumption in gasoline equivalent gallons (weighted average)	588.248	468.079	2	17953.2	3488	126
WHITE	Race of HH respondent: white	0.973108	0.161791	0	1	3614	0
WRKCNT2	Number of workers in HH - outliers removed	1.18647	0.834084	0	3	3536	78
WRKCOUNT	Number of workers in HH	1.22125	0.880385	0	5	3614	0
WRKTOSIZ	Number of workers to HH size ratio	0.540065	0.385218	0	1	3536	78
WTHHFIN	Final HH weight	3381.95	5592.67	5.4141	20126.6	3614	0

Appendix 1C: Correlation Matrix for Outliers

	DRIVER	SIZE	VEHCNT	ADULTS	WORKERS	COMMER	AUTO	PICKUP	OTHER	SUV	MGAS	MONTHS
DRIVER	1	0.2021	0.0796	0.5532	0.2893	0.0103	0.1486	0.0443	0.0834	0.0524	0.0874	-0.0063
SIZE	0.2021	1	0.0258	0.2412	0.1006	0.0058	0.0333	0.0179	0.0247	0.0102	0.0242	-0.0100
VEHCNT	0.0796	0.0258	1	0.0689	0.0702	0.0988	0.2482	0.2012	0.9711	0.0804	0.9138	0.0077
ADULTS	0.5532	0.2412	0.0689	1	0.2554	0.0133	0.1304	0.0297	0.0696	0.0295	0.0729	-0.0072
WORKERS	0.2893	0.1006	0.0702	0.2554	1	0.0138	0.1355	0.0446	0.0729	0.0408	0.0767	-0.0077
COMMER	0.0103	0.0058	0.0988	0.0133	0.0138	1	0.0206	0.1250	0.1055	0.0189	0.0815	-0.0020
AUTO	0.1486	0.0333	0.2482	0.1304	0.1355	0.0206	1	0.0109	0.2430	0.0021	0.2459	0.0049
PICKUP	0.0443	0.0179	0.2012	0.0297	0.0446	0.1250	0.0109	1	0.2140	0.0090	0.1844	0.0004
OTHER	0.0834	0.0247	0.9711	0.0696	0.0729	0.1055	0.2430	0.2140	1	0.0807	0.9406	0.0078
SUV	0.0524	0.0102	0.0804	0.0295	0.0408	0.0189	0.0021	0.0090	0.0807	1	0.0771	-0.0069
MGAS	0.0874	0.0242	0.9138	0.0729	0.0767	0.0815	0.2459	0.1844	0.9406	0.0771	1	0.0070
MONTHS	-0.0063	-0.0100	0.0077	-0.0072	-0.0077	-0.0020	0.0049	0.0004	0.0078	-0.0069	0.0070	1

Appendix 2A: Correlation Matrix – Nation

	LVMT	LEFFC
ALT	0.00652961	-0.0196809
AUTO	0.0170435	0.361729
BESTMILE	0.840959	0.170508
CENSUS_R	-0.00173489	-0.0652444
COMM	0.0733068	-0.0736965
CWEIADMP	0.23843	0.992424
DIESEL	0.0586149	-0.0744938
DRVRCNT	0.451708	0.0541926
DRVRCNT2	0.451708	0.0541926
ELECTRIC	-0.00280478	-1.02E-05
FED	0.0191656	0.0798409
GSCOST	-0.0388687	0.0319783
HBPPOPDN	-0.120206	0.0776112
HBRESDN	-0.128666	0.0677267
HH_HISP	0.0144852	0.0138957
HH_RACE	-0.00304211	0.0117762
HHFAMINC	0.367604	0.039653
HHSIZE	0.386526	0.0381108
HHSIZE2	0.386526	0.0381108
HHSTFIPS	0.0596424	-0.0220636
HHVEHCNT	0.503446	-0.00941045
HHVHCNT2	0.503446	-0.00941045
HIGH	-0.0463981	-0.00723091
HOMEOWN	0.0738169	-0.0883609
HOMETYPE	-0.0575259	0.0281462
HOUSEID	0.000238405	-0.00293049
HTEEMPDN	-0.129168	0.0599544
HTPPOPDN	-0.114096	0.0728678
HYBRID	0.037035	0.00706443
INCOME	0.363628	0.0381459
INCPCAP	-0.0123845	0.00783202
LAND	0.0401158	-0.00831564
LARGE	0.303802	-0.323056

	LVMT	**LEFFC**
LCONSUM	0.880795	0.0151754
LEFF	0.266238	1
LEFFC	0.266238	1
LEXP	0.879735	0.0167688
LGSCOST	-0.0389874	0.0327904
LIF_CYC	-0.183835	-0.0957744
LINCOME	0.35465	0.0381436
LMSASIZ	-0.0509081	0.0642487
LOW	-0.215898	-0.0242321
LVMT	1	0.266238
MED	-0.212568	-0.0251392
MGAS	0.0285069	0.0183708
MIDWEST	0.0186685	0.0297474
MINOR	-0.00639938	0.00987763
MSACAT	0.0345967	-0.0846425
MSASIZ	-0.0219729	0.0685371
MSASIZE	0.037807	0.0024543
N_AUTO	0.142264	0.370412
N_AUTO2	0.142264	0.370412
N_DIESEL	0.0602763	-0.0756117
N_ELECTR	-0.00280478	-1.02E-05
N_HYBRID	0.0393004	0.00443442
N_MGAS	0.491708	0.00321861
N_MGAS2	0.491708	0.00321861
N_NGAS	0.00732839	-0.0186922
N_OTHER	0.537075	-0.0273405
N_OTHER2	0.537075	-0.0273405
N_PICKUP	0.236566	-0.268755
N_PICUP2	0.236566	-0.268755
N_SUV	0.221604	-0.2299
N_SUV2	0.221604	-0.2299
N_VEHCO2	0.0717486	-0.0712938
N_VEHCOM	0.0717486	-0.0712938
NGAS	0.00702001	-0.0198966
NORTH	-0.0245766	0.0824867
NOSUB	-0.490575	0.0412526
NUMADLT	0.376269	0.0415397

	LVMT	**LEFFC**
NUMADLT2	0.376269	0.0415397
ONEADULT	-0.370162	0.0197287
OTHER	0.471851	-0.0589374
PICKUP	0.22574	-0.263428
RAIL	-0.0141037	0.0789467
RATEXP	0.200319	-0.0413169
RATEXPC	0.0243593	-0.0156705
SGSTOTCS	0.648858	-0.0747384
SGSYRGAL	0.65533	-0.0745631
SOUTH	0.0281933	-0.0859433
STATE	-0.0427599	0.0677561
SUV	0.210791	-0.223322
TAX	-0.0427599	0.0677561
TPM	-0.224134	-0.582987
URB	-0.127747	0.0639583
URBAN	0.12841	-0.0689851
URBANSIZ	0.0901438	-0.0288292
URBRUR	0.127747	-0.0639583
VEHAGE	0.046174	-0.117243
VEHOWNM2	-0.0224695	-0.124296
VEHOWNMO	-0.0224695	-0.124296
VHIGH	0.31586	0.0380659
WEIADMPG	0.23843	0.992424
WEST	-0.0313203	0.0169286
WGSTOTCS	0.625414	-0.0730565
WGSYRGAL	0.636794	-0.0728082
WHITE	0.00143535	-0.0418313
WRKCNT2	0.428051	0.136293
WRKCOUNT	0.428051	0.136293
WRKTOSIZ	0.267622	0.133884
WTHHFIN	0.0204867	0.0853441

Appendix 2B: Correlation Matrix – Iowa

	LVMT	LEFFC
ALT	0.007425	-0.05113
AUTO	0.012722	0.350547
BESTMILE	0.84689	0.152017
CENSUS_R	0	0
COMM	0.098445	-0.08065
CWEIADMP	0.228311	0.992229
DIESEL	0.109081	-0.06411
DRVRCNT	0.432676	0.058527
DRVRCNT2	0.432676	0.058527
ELECTRIC	0	0
FED	0	0
GSCOST	0.053084	-0.05335
HBPPOPDN	-0.2089	0.024866
HBRESDN	-0.22623	0.00851
HH_HISP	-0.03992	0.031409
HH_RACE	-0.00817	-0.01997
HHFAMINC	0.37521	0.062444
HHSIZE	0.373354	0.051184
HHSIZE2	0.373354	0.051184
HHSTFIPS	0.118221	-0.04485
HHVEHCNT	0.518317	0.034719
HHVHCNT2	0.518317	0.034719
HIGH	-0.04863	0.011363
HOMEOWN	0.031017	-0.00217
HOMETYPE	-0.03696	0.014654
HOUSEID	-0.00184	0.029638
HTEEMPDN	-0.12966	0.052832
HTPPOPDN	-0.18274	0.024319
HYBRID	0.007777	-0.09928
INCOME	0.370274	0.062314
INCPCAP	0.019285	-0.00675
LAND	0.043444	-0.05823
LARGE	0.314915	-0.30455

	LVMT	**LEFFC**
LCONSUM	0.865624	0.029133
LEFF	0.257467	1
LEFFC	0.257467	1
LEXP	0.865239	0.027996
LGSCOST	0.051641	-0.04939
LIF_CYC	-0.2015	-0.05499
LINCOME	0.372773	0.056193
LMSASIZ	-0.11871	0.093793
LOW	-0.21555	-0.03277
LVMT	1	0.257467
MED	-0.20738	-0.04126
MGAS	0	0
MIDWEST	0	0
MINOR	-0.01422	-0.00472
MSACAT	0.123551	-0.09464
MSASIZ	-0.02263	0.049706
MSASIZE	0.13295	-0.09411
N_AUTO	0.158257	0.364158
N_AUTO2	0.158257	0.364158
N_DIESEL	0.113517	-0.07155
N_ELECTR	0	0
N_HYBRID	0.00901	-0.09382
N_MGAS	0.503417	0.049283
N_MGAS2	0.503417	0.049283
N_NGAS	0.007425	-0.05113
N_OTHER	0.550683	0.017815
N_OTHER2	0.550683	0.017815
N_PICKUP	0.253586	-0.27076
N_PICUP2	0.253586	-0.27076
N_SUV	0.182507	-0.17839
N_SUV2	0.182507	-0.17839
N_VEHCO2	0.075583	-0.08644
N_VEHCOM	0.075583	-0.08644
NGAS	0.007425	-0.05113
NORTH	0	0
NOSUB	-0.45692	0.015434
NUMADLT	0.352594	0.055287

	LVMT	**LEFFC**
NUMADLT2	0.352594	0.055287
ONEADULT	-0.35689	-0.02863
OTHER	0.446063	-0.0297
PICKUP	0.251278	-0.24767
RAIL	0	0
RATEXP	0.287764	-0.04111
RATEXPC	0.058805	-0.02438
SGSTOTCS	0.644807	-0.06503
SGSYRGAL	0.651986	-0.05914
SOUTH	0	0
STATE	0.118221	-0.04485
SUV	0.1744	-0.1638
TAX	0.118221	-0.04485
TPM	-0.28638	-0.99159
URB	-0.21853	0.06387
URBAN	0.223793	-0.07319
URBANSIZ	0.193951	-0.07932
URBRUR	0.218528	-0.06387
VEHAGE	0.028767	-0.09448
VEHOWNM2	-0.10356	-0.10637
VEHOWNMO	-0.10356	-0.10637
VHIGH	0.314384	0.03987
WEIADMPG	0.228311	0.992229
WEST	0	0
WGSTOTCS	0.610613	-0.08392
WGSYRGAL	0.618532	-0.07767
WHITE	0.016031	0.003718
WRKCNT2	0.443538	0.095147
WRKCOUNT	0.443538	0.095147
WRKTOSIZ	0.264361	0.063985
WTHHFIN	0.123676	-0.04692

Appendix 3A: ANOVA – Average Vehicle Fuel Efficiency at the HH Level

Summary of fit	
R squared	5.77e-6
R squared adjusted	-1.57e-6
Root mean square error	0.956606
Mean of response	9.604053
Observations	136176

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F-ratio
Model	1	0.72	0.719798	0.7866
Error	136174	124612.10	0.915095	Prob>F
C. Total	136175	124612.82		0.3751

Parameter Estimates				
Term	Estimate	Std. Error	t-Ratio	Prob>|t|
Intercept	9.6110473	0.008301	1157.8	<0.0001
Region (Iowa)	0.0073623	0.008301	0.89	0.3751

Effect Tests					
Source	Nparm	DF	Sum of Squares	F-ratio	Prob>F
Region	1	1	0.71979848	0.7866	0.3751

Least Squares Means Table			
Level	Least Sq. Mean	Std. Error	Mean
Iowa	9.6184096	0.04639360	9.61841
Nation	9.6036850	0.00262531	9.60368

Appendix 3B: ANOVA – Vehicle-miles traveled at the HH Level

Summary of fit	
R squared	1.955e-5
R squared adjusted	1.221e-5
Root mean square error	0.196088
Mean of response	3.012617
Observations	136176

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F-ratio
Model	1	0.1024	0.102364	2.6622
Error	136174	5235.9494	0.038450	Prob>F
C. Total	136175	5236.0518		0.1028

Parameter Estimates				
Term	Estimate	Std. Error	t-Ratio	Prob>|t|
Intercept	3.0099798	0.001702	1768.9	<0.0001
Region (Iowa)	-0.002776	0.001702	-1.63	0.1028

Effect Tests					
Source	Nparm	DF	Sum of Squares	F-ratio	Prob>F
Region	1	1	0.10236362	2.662	0.1028

Least Squares Means Table			
Level	Least Sq. Mean	Std. Error	Mean
Iowa	3.0072034	0.00336041	3.00720
Nation	3.0127562	0.00053814	3.01276

Appendix 4A: LIMDEP Output - Nation

```
**********************************************************************
 * NOTE: Deleted  26897 observations with missing data. N is now 116045 *
 **********************************************************************

Criterion function is max(abs(%chg in b(i))).
Iteration    0, 3SLS          =    1.000000
Iteration    1, 3SLS          =    .4028822
Iteration    2, 3SLS          =    .2239772E-01

+----------------------------------------------------+
| Estimates for equation: LEFFC                      |
| InstVar/GLS least squares regression               |
| Model was estimated Mar 26, 2012 at 00:31:50PM     |
| LHS=LEFFC    Mean                 =   3.028888     |
|              Standard deviation   =   .2037340     |
| WTS=WTHHFIN  Number of observs.   =     116045     |
| Model size   Parameters           =         14     |
|              Degrees of freedom   =     116031     |
| Residuals    Sum of squares       =   3025.929     |
|              Standard error of e  =   .1614888     |
| Fit          R-squared            =   .3717079     |
|              Adjusted R-squared   =   .3716375     |
| Model test   F[ 13,116031] (prob) =5280.45 (.0000) |
| Diagnostic   Log likelihood       =   46933.40     |
|              Restricted(b=0)      =   19960.44     |
|              Chi-sq [ 13]  (prob) =******* (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -3.646519     |
|              Akaike Info. Criter. =  -3.646519     |
| Not using OLS or no constant. Rsqd & F may be < 0. |
| Durbin-Watson 2.016 Autocorrelation =       -.0078 |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|    2.13884215      .02600874    82.236   .0000
 LARGE   |    -.18207059      .00120930  -150.559   .0000    .56004539
 LVMT    |     .10306875      .00258436    39.882   .0000   9.57927837
 WRKTOSIZ|     .03603088      .00165027    21.833   .0000    .51453897
 VEHOWNM2|    -.00015545    .154994D-04   -10.030   .0000   51.9879655
 VEHAGE  |    -.00405495      .00012891   -31.455   .0000   7.59312450
 HOMEOWN |    -.01302756      .00098946   -13.166   .0000    .71514646
 RAIL    |     .01215820      .00115457    10.531   .0000    .25774755
 TAX     |     .03610489      .00453457     7.962   .0000    .49440707
 URB     |     .01803331      .00122111    14.768   .0000    .75854016
 NOSUB   |     .01048025      .00237740     4.408   .0000    .33757156
 ONEADULT|     .01186838      .00134532     8.822   .0000    .28014399
 LAND    |    -.00875969      .00144130    -6.078   .0000    .87309374
 HYBRID  |     .01387375      .00226353     6.129   .0000    .03318253
```

```
+----------------------------------------------------+
| Estimates for equation: LVMT                       |
| InstVar/GLS least squares regression               |
| Model was estimated Mar 26, 2012 at 00:31:50PM     |
| LHS=LVMT     Mean                 =   9.579278     |
|              Standard deviation   =   .9751794     |
| WTS=WTHHFIN  Number of observs.   =     116045     |
| Model size   Parameters           =         13     |
|              Degrees of freedom   =     116032     |
| Residuals    Sum of squares       =   51284.68     |
|              Standard error of e  =   .6648213     |
| Fit          R-squared            =   .5352231     |
|              Adjusted R-squared   =   .5351750     |
| Model test   F[ 12,116032] (prob) =******* (.0000) |
| Diagnostic   Log likelihood       =  -117280.4     |
|              Restricted(b=0)      =  -161743.6     |
|              Chi-sq [ 12]  (prob) =******* (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -.8163619     |
|              Akaike Info. Criter. =  -.8163619     |
| Not using OLS or no constant. Rsqd & F may be < 0. |
| Durbin-Watson 1.988 Autocorrelation =        .0058 |
+----------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Constant	.83462276	.25658337	3.253	.0011	
LEFFC	2.29116621	.08954689	25.586	.0000	3.02888814
NOSUB	-.45289802	.00659889	-68.632	.0000	.33757156
DRVRCNT2	.21832916	.00407676	53.555	.0000	1.82601549
LINCOME	.14009963	.00326817	42.868	.0000	10.6735466
WRKTOSIZ	.13345160	.00818681	16.301	.0000	.51453897
LARGE	.55197758	.01510796	36.536	.0000	.56004539
URB	-.14064465	.00504672	-27.869	.0000	.75854016
LMSASIZ	-.00376332	.00033439	-11.254	.0000	11.2665444
TAX	-.18728957	.01871756	-10.006	.0000	.49440707
LAND	.01996015	.00601547	3.318	.0009	.87309374
VEHOWNM2	-.00147321	.807594D-04	-18.242	.0000	51.9879655
RAIL	-.05095087	.00497779	-10.236	.0000	.25774755

Appendix 4B: LIMDEP Output - Iowa

```
*********************************************************************
 * NOTE: Deleted    620 observations with missing data. N is now   2994 *
 *********************************************************************

Criterion function is max(abs(%chg in b(i))).
Iteration    0, 3SLS          =    1.000000
Iteration    1, 3SLS          =    .5752372
Iteration    2, 3SLS          =    .2661053E-02

+---------------------------------------------------+
| Estimates for equation: LEFFC                     |
| InstVar/GLS least squares regression              |
| Model was estimated Mar 26, 2012 at 00:37:54PM    |
| LHS=LEFFC    Mean                 =   3.012885    |
|              Standard deviation   =   .1807634    |
| WTS=WTHHFIN  Number of observs.   =       2994    |
| Model size   Parameters           =          8    |
|              Degrees of freedom   =       2986    |
| Residuals    Sum of squares       =   69.16858    |
|              Standard error of e  =   .1521982    |
| Fit          R-squared            =   .2908420    |
|              Adjusted R-squared   =   .2891795    |
| Model test   F[  7,  2986] (prob) = 174.95 (.0000) |
| Diagnostic   Log likelihood       =   1392.123    |
|              Restricted(b=0)      =   873.6330    |
|              Chi-sq [   7]  (prob) =1036.98 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -3.762475    |
|              Akaike Info. Criter. =  -3.762475    |
| Not using OLS or no constant. Rsqd & F may be < 0. |
| Durbin-Watson 2.010 Autocorrelation =      -.0049 |
+---------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|    2.42730851       .06737795    36.025   .0000
 LARGE   |    -.10333807       .00523823   -19.728   .0000    .74348697
 LVMT    |     .07158181       .00713136    10.038   .0000   9.61919407
 VEHOWNM2|    -.00040209     .781071D-04    -5.148   .0000   54.6396329
 HYBRID  |    -.03528649       .01032310    -3.418   .0006    .06145625
 WRKTOSIZ|     .03369303       .00816055     4.129   .0000    .48483634
 VEHAGE  |    -.00387050       .00070489    -5.491   .0000   7.67586705
 LMSASIZ |     .00165603       .00047221     3.507   .0005   6.86001176
```

```
+----------------------------------------------------+
| Estimates for equation: LVMT                       |
| InstVar/GLS least squares regression               |
| Model was estimated Mar 26, 2012 at 00:37:54PM     |
| LHS=LVMT     Mean                 =   9.619194     |
|              Standard deviation   =   .9031439     |
| WTS=WTHHFIN  Number of observs.   =       2994     |
| Model size   Parameters           =          9     |
|              Degrees of freedom   =       2985     |
| Residuals    Sum of squares       =   1244.585     |
|              Standard error of e  =   .6457137     |
| Fit          R-squared            =   .4886583     |
|              Adjusted R-squared   =   .4872878     |
| Model test   F[  8,  2985] (prob) = 356.57 (.0000) |
| Diagnostic   Log likelihood       =  -2934.223     |
|              Restricted(b=0)      =  -3942.793     |
|              Chi-sq [  8]  (prob) =2017.14 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -.8717965     |
|              Akaike Info. Criter. =  -.8717965     |
| Not using OLS or no constant. Rsqd & F may be < 0. |
| Durbin-Watson 2.023 Autocorrelation =      -.0114 |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|    -.64816904      1.03601389     -.626   .5316
 LEFFC   |    2.75780688       .37026389     7.448   .0000   3.01288524
 NOSUB   |    -.50901388       .03856646   -13.198   .0000    .21810287
 ONEADULT|    -.20221465       .03335638    -6.062   .0000    .22745491
 URB     |    -.14846501       .02326129    -6.382   .0000    .63694055
 LINCOME |     .17812482       .02108763     8.447   .0000   10.7152408
 LMSASIZ |    -.00898180       .00213816    -4.201   .0000   6.86001176
 LARGE   |     .38057117       .03939932     9.659   .0000    .74348697
 WRKTOSIZ|     .16500042       .03839283     4.298   .0000    .48483634
```

Biographical Sketch

Eirini Kastrouni was born June 19th, 1986 in Athens, Greece. She received the Bachelor of Science in Civil Engineering from the National Technical University of Athens in 2010. She joined Iowa State University in August 2010 to pursue a Master of Science in Civil Engineering, with an emphasis in Transportation Engineering. During her graduate studies, she has served as a Research Assistant in the Institute for Transportation, working in the NSF-funded research project NETSCORE21. In recognition of her academic excellence, she has been invited to join Tau Beta Pi Engineering Honor Society, and Honor Society of Phi Kappa Phi. She is also an Executive Fellow for the 2012 IRF Road Scholar Program. Regarding her leadership activity, she has been elected cabinet member for the Transportation Student Association from the 2nd semester of her graduate studies, and she has also been elected Vice President for the 2012 IRF Class.