ESSAYS ON THE ECONOMICS OF MOTOR VEHICLE ENERGY EFFICIENCY

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Tingmingke Lu

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Stephen Martin, Chair

Krannert School of Management

Dr. Trevor S. Gallen

Krannert School of Management

Dr. Joseph Mazur

Krannert School of Management

Dr. Justin L. Tobias

Krannert School of Management

Approved by:

Dr. Brian Roberson

Head of Graduate Program in Economics

This dissertation is dedicated to my mom and my wife, for their support and encouragement.

ACKNOWLEDGMENTS

I want to thank my advisor, Stephen Martin, for all his patience, guidance, and good advice throughout my time at Purdue. I am grateful to the members of my committee, Trevor Gallen, Joe Mazur, and Justin Tobias, for their invaluable suggestions and support. I also would like to thank Tim Cason for offering help and advice during the job market process. Thanks to Emmanuel Dechenaux for leading me down the path of scholarship. Thanks to Justin Barnette and Jooyoun Park for always being supportive and having faith in me. This research has benefited from conversations with Federico Rossi and Anson Soderbery. I am grateful for their comments and help.

I wish to thank my officemates for all the help and inspiring discussions. I am also indebted to my family and friends for their companionship. Finally, special thanks to the Economics Department at Purdue University for providing me with the opportunity and training.

TABLE OF CONTENTS

Page

LIST OF TABLES . . . vii

LIST OF FIGURES . . . viii

ABSTRACT . . . ix

1 RESPONSE OF NEW CAR BUYERS TO ALTERNATIVE ENERGY POLICIES: THE ROLE OF VEHICLE USE HETEROGENEITY . . . 1
- 1.1 Introduction . . . 1
- 1.2 Related Literature . . . 4
- 1.3 An Empirical Model of New Car Demand . . . 6
  - 1.3.1 Model Formulation . . . 7
  - 1.3.2 Individual Choice Probability and Market Share . . . 10
  - 1.3.3 Heterogeneous Consumer Response to Alternative Policies . . . 11
- 1.4 Data . . . 13
  - 1.4.1 Vehicle Choices . . . 13
  - 1.4.2 New Car Markets . . . 14
  - 1.4.3 Prices . . . 15
  - 1.4.4 Empirical Distribution of Annual Vehicle Miles Traveled . . . 16
- 1.5 Empirical Implementation . . . 17
  - 1.5.1 Specification of the Taste Parameters . . . 17
  - 1.5.2 Identification and the GMM Estimator . . . 19
- 1.6 Results . . . 21
  - 1.6.1 Parameter Estimates . . . 22
  - 1.6.2 Consumer Valuation of Expected Gas Cost Savings . . . 24
- 1.7 Policy Simulations . . . 26
  - 1.7.1 Perfectly Inelastic Demand for Driving . . . 26
  - 1.7.2 Elastic Demand for Vehicle Miles Traveled . . . 29
- 1.8 Concluding Remarks . . . 30

2 IDENTIFYING THE ELASTICITY OF DRIVING: A FACTOR MODEL APPROACH AND IMPLICATIONS FOR USAGE-BASED TRANSPORTATION POLICIES . . . 39
- 2.1 Introduction . . . 39
- 2.2 Related Literature . . . 41
- 2.3 Data . . . 43
  - 2.3.1 Data Sources . . . 43
  - 2.3.2 Data Description . . . 45

Page

2.4 Empirical Approach 47

2.4.1 Model Specification 47

2.4.2 Identification 48

2.4.3 Empirical Implementation 50

2.5 Results 52

2.6 Concluding Remarks 55

REFERENCES 62

A APPENDIX: RESPONSE OF NEW CAR BUYERS TO ALTERNATIVE ENERGY POLICIES 67

A.1 Consumer Response to Policy Changes 67

A.2 Data 69

A.3 Estimation Details and Numerical Considerations 74

A.4 Revenue-equivalent Product Taxes 75

A.5 Elastic Demand for Driving 76

B APPENDIX: IDENTIFYING THE ELASTICITY OF DRIVING 77

B.1 Instrumental Variable for Massachusetts Motor Fuel Prices 77

VITA 78

LIST OF TABLES

Table Page

1.1 New Passenger Vehicle Registration Records 33
1.2 Demand Sample and Mileage Sample 34
1.3 Vehicle Attributes 35
1.4 Parameter Estimates from Alternative Models 36
1.5 Consumer Valuation of Expected Gas Cost Savings 37
1.6 Expected Average Annual Gasoline Consumption 38
2.1 Summary Statistics by Regions 59
2.2 Regression Results for Variable Driving Cost 60
2.3 Regression Results for Retail Gas Price 61
A.1 New Vehicle Registration Counts in 2011 70
A.2 Vehicle Counts by Region and by Month, Demand Sample 72
A.3 Vehicle Counts by Region and by Month, Mileage Sample 73
A.4 The Elasticity of Driving in Mileage Sample 76
B.1 First-stage Regression Results, Two-way Fixed Effects Case 77

LIST OF FIGURES

Figure Page

1.1 Monthly Average Price of Midgrade Gasoline 32

2.1 Fleet Fuel Efficiency and Massachusetts Real Fuel Price 57

2.2 Daily Miles Traveled Per Vehicle and Variable Driving Cost 58

A.1 Metropolitan Planning Organization Regions 71

ABSTRACT

Lu, Tingmingke Ph.D., Purdue University, August 2019. Essays on the Economics of Motor Vehicle Energy Efficiency. Major Professor: Stephen Martin.

The purpose of this dissertation is to study the effectiveness of public policies in generating fuel savings and emissions reductions. I focus on applying various empirical methods to analyze consumer responses to policy changes on both extensive and intensive margins. This dissertation consists of two chapters.

In the first chapter, I compare the effectiveness of fuel taxes and product taxes on reducing gasoline consumption of new car buyers. I employ a unified data source for vehicle choice and subsequent vehicle use to estimate a random effects logit demand model that explicitly accounts for vehicle use heterogeneity. My demand estimation suggests that new car buyers fully value the fuel-saving benefits from improved vehicle fuel efficiency when they initially purchase their cars. My policy simulations indicate that high-mileage drivers are more responsive to a change in fuel taxes than to a change in product taxes, even as low-mileage drivers are more responsive to product taxes. By capturing such heterogeneous consumer response to policies, I show that a counterfactual increase of the fuel tax is more effective than a revenue-equivalent product tax in reducing the total gasoline consumption of new car buyers. Further, when accounting for its effects on consumer response on both extensive and intensive margins, a change in fuel taxes has a clear advantage over a change in product taxes in reducing the consumption of gasoline even when the magnitude of tax increase is small. More importantly, a model not accounting for vehicle use heterogeneity understates the fuel saving effects of both policies and misleads us about the relative effectiveness when comparing different policies.

The second chapter explores how changes in the marginal cost of driving affect consumers decisions about passenger vehicle utilization, as measured by average daily miles traveled per vehicle. This intensive margin of consumer response has important implications for the effectiveness of usage-based policies, such as the fuel tax and the mileage tax, that designed to address externalities of driving. I estimate the elasticity of driving with respect to fuel cost per mile using a large panel data that covers 351 towns and cities in Massachusetts over 24 quarters. While most researchers in this literature apply fixed effects estimators to examine the elasticity of driving, I use a factor model econometric setup to account for unobserved common factors and regional heterogeneity. Residual diagnostics confirm that the factor model setup does a better job of removing the cross-section dependence than fixed effects estimators do. Given low consumer responsiveness to changes in the marginal cost of driving engendered by current usage-based policies, rights-based approaches like congestion charges might be better alternatives to influence vehicle utilization and vehicle ownership.

1. RESPONSE OF NEW CAR BUYERS TO ALTERNATIVE ENERGY POLICIES: THE ROLE OF VEHICLE USE HETEROGENEITY

1.1 Introduction

In the United States, gasoline consumption by passenger cars and light-duty trucks accounts for about 59 percent of the carbon emissions attributed to transportation activities, which collectively make up the most significant share of U.S. carbon emissions from fossil fuel consumption among all end-use economic sectors (EPA, 2018). At the local level, the ambient air pollution caused by automobile emissions has been found to affect infant health and contribute to the high rate of acute asthma attacks among young children in some urban areas (Knittel et al., 2016; Simeonova et al., 2018). Growing concerns about energy overuse, climate change, and the impact of local air pollution on public health have raised the interest in designing effective public policies to reduce motor fuel consumption in the passenger transportation sector.

This paper investigates the effects of two policy instruments on gasoline conservation. In particular, I compare the effectiveness of energy taxes and product taxes in reducing total gasoline consumption. An energy tax could be a fuel tax, a carbon tax, or other policies that change the retail energy price. A product tax refers to any policy that alters the relative purchase prices of energy-intensive durable goods with different energy efficiency levels, such as fees imposed on inefficient vehicles and rebates provided for efficient vehicles according to their per-mile fuel consumption.

I show that the fuel tax outperforms the product tax in reducing gasoline consumption of new car buyers because high-mileage drivers are more responsive to a change in fuel taxes than to a change in product taxes, even as low-mileage drivers are more responsive to product taxes. I argue that accounting for vehicle use heterogeneity is critical because it helps address the endogenous sorting of low-mileage drivers into less fuel-efficient car models in demand estimation and it more accurately captures the heterogeneous consumer response to alternative energy policies. Moreover, models not accounting for vehicle use heterogeneity understate the fuel saving effects of both policies and mislead us about the relative effectiveness when comparing different policies.

I model the effects of fuel taxes and product taxes on consumers' vehicle choices as in Grigolon et al. (2018). When a product tax is applied, vehicles with higher fuel economy ratings become cheaper, *ceteris paribus*, relative to other models, which incentivizes consumers to choose fuel-efficient vehicles. A fuel tax influences consumers' vehicle purchase decisions by augmenting the difference in the fuel cost between vehicle models with different fuel economy ratings. The success of this mechanism depends on how many miles a consumer drives and to what extent this consumer values the future gas cost savings arising from improved vehicle fuel efficiency.

To characterize consumer heterogeneity in vehicle use intensity and preference over vehicle attributes, I assemble the vehicle ownership and inspection history for all new passenger cars sold in Massachusetts in 2011 from a dataset that contains information about every vehicle registered in Massachusetts from 2009 to 2014. I calculate the annual miles traveled of a car using the difference between odometer readings from the vehicle inspection logs.

My empirical analysis proceeds in two steps. First, I estimate a random effects logit demand model in the style of Berry et al. (1995). In particular, I apply the discrete choice model to define the probability of a consumer buying a specific new car model as a function of the car's attributes (including price) and the present discounted value of the car's lifetime gas cost. To incorporate vehicle use heterogeneity, I let the new car buyer's expected annual vehicle miles traveled and the car's fuel efficiency jointly determine the car's lifetime gas cost. Variations in car attributes and expected lifetime gas cost allow me to identify flexible substitution patterns for new car buyers. I model unobserved car attributes as random effects, and I employ price shifters as instrumental variables to address price endogeneity. In the second step, I conduct a counterfactual analysis to understand the effects of alternative energy policies on gasoline conservation. I compare the amount of fuel savings resulting from applying a fuel tax and a product tax on the per-mile fuel consumption of a car model, holding revenues equivalent.

My estimates show that new car buyers fully value the gas cost savings generated from improved vehicle fuel efficiency. They are willing to pay an extra one dollar in the up-front purchase price of the car for a one-dollar reduction in discounted future gas cost. My calculation is closer to the full-valuation results reported by Busse et al. (2013) and Grigolon et al. (2018) than to the partial-valuation results obtained by Allcott and Wozny (2014) and Leard et al. (2017).

It is crucial to understand whether consumers are myopic in their preferences about vehicle fuel efficiency. For instance, if consumers are fully rational and willing to pay for improved fuel efficiency, product taxes would not be necessary. However, if consumers are boundedly rational or have a low willingness to pay for improved fuel efficiency, product taxes will serve the same purpose as subsidies in creating price differences and stimulating more consumers to buy fuel-efficient car models than they otherwise would.

By accurately capturing heterogeneous consumer response to alternative policies, my counterfactual analysis shows that drivers in the top half of the annual vehicle miles traveled distribution are more responsive to the fuel tax than to the product tax. Therefore, the fuel tax incentivizes high-mileage drivers more to choose fuel-efficient car models and to substitute away from conventional gasoline-powered cars toward hybrid and electric vehicles.

In addition, I relax the assumption of perfectly inelastic demand for driving in my policy simulations. I show that when accounting for consumer responses on both extensive and intensive margins, a change in fuel taxes has a clear advantage over a change in product taxes in reducing gasoline consumption of new car buyers even when the magnitude of tax increase is small.

The rest of the paper is laid out as follows. The next section reviews the relevant literature. Section 3 develops the demand model and explains how consumers respond differently to the change of alternative policies when heterogeneity in annual mileage is considered. Section 4 describes the data. In section 5, I introduce the empirical framework and state the identification strategy. Section 6 presents the results from demand estimation and discusses the implied consumer valuation of expected future gas cost savings. Section 7 carries out the counterfactual analysis. Section 8 concludes.

1.2 Related Literature

This paper is related to a growing literature investigating whether there is underinvestment in vehicle fuel efficiency by consumers. I contribute to this literature by explicitly modeling the consumer heterogeneity inherent in vehicle use intensity using the mileage distribution of new car buyers constructed from high-quality vehicle level microdata, and finding that accounting for vehicle use heterogeneity is crucial. Bento et al. (2012) provide simulation evidence to show that ignoring consumer heterogeneity when estimating the willingness to pay for discounted future energy cost savings will result in a sorting bias, which may mistakenly lead to the conclusion of consumer undervaluation.

Using monthly vehicle sales data from 1999 to 2008, Busse et al. (2013) study how changes in gasoline price affect the equilibrium prices and quantities sold of new and used cars with different fuel economy ratings. They conclude that consumers are not myopic because the calculated implicit discount rates of consumers are similar to the interest rates paid by car buyers who borrow.[1] Employing similar reduced-form price and quantity regressions, Allcott and Wozny (2014) find used car buyers

[1] In a recent working paper, Leard et al. (2017) apply the same approach to households new vehicles purchase data between 2009 and 2014. They find that consumers are only willing to pay about 54 cents for one dollar of the expected gas cost savings but have high willingness to pay for vehicle performances, which implies approximately zero net private consumer benefit from tightened fuel economy standards.

are indifferent between 76 cents in the vehicle purchase price and one dollar in the present discounted value of future gas cost.

In this paper, I take advantage of detailed microdata. I use information about observed vehicle choices and subsequent vehicle use to capture consumer heterogeneity in annual mileage. In contrast, most studies in this literature apply information from the National Household Travel Survey (NHTS) to calculate the annual mileage by vehicle class and by vehicle age at the national level.

In addition, while estimates in Busse et al. (2013) and Allcott and Wozny (2014) mainly rely on fuel price differences and hence induced fuel cost changes, I focus on fuel cost variation generated by heterogeneity in vehicle use intensity. Gillingham et al. (2015) find substantial heterogeneity in consumer response to gasoline prices by vehicle fuel economy quantiles. This implies consumers may value a decrease in fuel price and an increase in vehicle fuel efficiency differently because the former can be seen as a relatively short-term gain in comparison to the latter.

Grigolon et al. (2018) provide the first study treating consumers' vehicle use heterogeneity and their valuation of expected fuel cost savings in a unified empirical framework. They apply a discrete distribution of the annual miles traveled obtained from the 2007 U.K. National Travel Survey and assign consumers to fourteen mileage types when estimating the new car demand in seven European countries over fourteen years. The authors find a modest undervaluation of expected fuel cost savings among consumers. My paper employs the empirical mileage distributions constructed from the inspection logs of vehicles sold in separate new car markets. The observed annual vehicle miles traveled during years after the new car purchase provide a better approximation of the car's lifetime mileage.

This paper is also related to a strand of literature that asks how to set and evaluate energy policies in the presence of heterogeneous market failures.[2] Allcott et al. (2014) point out the first best policy in this case would deliver heterogeneous corrections to consumers' decision utility. A consumer-specific product subsidy may serve this

[2]This literature includes, for instance, Innes (1996), Fullerton and West (2002), and Fullerton and West (2010).

purpose if it is designed to address the individual level bias of each consumer. In the absence of such first best policy, I explore the empirical distribution of vehicle miles traveled and show that high-mileage drivers are more responsive to the fuel tax than to the product tax. Hence, the fuel tax is more effective in generating fuel savings and emissions reductions relative to a uniform product tax.

Through policy counterfactuals using European data, Grigolon et al. (2018) document that even if the demand for vehicle miles traveled is perfectly inelastic and consumers do not fully value expected future fuel cost savings, a fuel tax can still be more effective than a product tax if consumers are heterogeneous in vehicle use. In the United States, the retail gasoline price is about 50 percent lower than that in European countries as of 2011, while the average annual miles traveled per vehicle is about 20 percent higher. Although there are significant differences between Europe and the U.S. in terms of gas prices and miles traveled, I show that a fuel tax is preferred over product taxes for gasoline conservation when there is no systematic undervaluation of vehicle fuel efficiency among consumers. This is because the fuel tax better targets high-mileage drivers. Consequently, the high-mileage drivers are more inclined to substitute toward fuel-efficient car models.

1.3 An Empirical Model of New Car Demand

In this section, I present a random coefficients discrete choice model of the demand for new conventional passenger cars in the style of Berry et al. (1995) and Petrin (2002). I let the present discounted value of a car's lifetime gas cost enter the consumer's decision of which new car to purchase as in Grigolon et al. (2018). Since consumers may differ in expected vehicle miles traveled and hence in their expected gas costs, I incorporate the empirical distribution of annual mileage when constructing the expected gas cost variable.

I employ the sample of new car buyers and estimate the conditional choice of buying a new passenger car directly. In a complete vehicle choice model, the consumer can choose to buy a new vehicle, buy a used vehicle, continue using her current vehicle, or not own any vehicle. To include all these options, I need to weight the sample of new car buyers, so it is consistent with the general population. However, I do not have enough information to correctly infer the conditional density of tastes, and hence, the weights of new vehicle buyers, from the population density. Incorrectly applied weights will lead me to inconsistent estimates of tastes that affect both which new car the consumer chooses and whether the consumer chooses other alternative options such as not holding a car (Train and Winston, 2007).[3]

My data comes from a regional market (Massachusetts) and covers six months.[4] Thus, I model and analyze the effects of alternative energy policies on the demand system in the short run. In my model, the product offering is exogenous to policy changes, so the supply side is not specified or estimated. I drop diesel-powered vehicles because the number of new diesel cars sold in a single month is small.[5] I define the inside good as gasoline-powered passenger cars excluding SUVs. The outside good contains both hybrid and electric cars.

1.3.1 Model Formulation

There are T markets with I_t potential buyers of new passenger cars in each market t. The conditional indirect utility function of consumer i for car v is

$$u_{ivt} = \alpha_i \; x_v + \beta_i \; (p_v + \gamma \; G_{ivt}) + \zeta_{vt} + \epsilon_{ivt}, \tag{1.1}$$

[3]A further difficulty with considering the purchase of used vehicles is that used car prices are not available because I do not observe the used car transactions from data.

[4]Records of almost all vehicles registered in Massachusetts from 2009 to 2014 are available. I form my sample to avoid the automotive industry crisis of 2008 - 2010. I choose to use the model year overlapping months during which both new model year 2011 and 2012 cars are for sale because these observations have the longest inspection histories within the feasible time frame for me to trace mileage records.

[5]Among all passenger vehicles registered in Massachusetts as of 2017, 95.56 percent are powered by gasoline, 3.37 percent are hybrid and electric vehicles, and only 1.07 percent are powered by diesel fuel. (Autoalliance.org, State Facts: Autos drive Massachusetts forward, 2017).

in which x_v is a vector of observed car attributes, p_v is the purchase price of car v, and G_{ivt} is the present discounted value of expected future gas cost associated with consumer i choosing to buy car v in market t. This indirect utility does not include the income of consumer i because the income is common to all options in her choice set and drops out of the equation eventually (Nevo, 2000; Train, 2009). Coefficients included in vector α_i identify the individual-specific valuation of observed car attributes, while β_i is the individual-specific price sensitivity. ζ_{vt} represents the unobserved (to the econometrician) vehicle attributes of model v in market t. The unobserved component of the utility function ϵ_{ivt} follows the Type I Extreme Value Distribution, and it is independent and identically distributed over all consumers i, cars v, and markets t. I normalize the utility of the outside good to zero so that my estimates are in terms of the difference between the utility of purchasing a specific car model v and the utility of choosing to buy the outside good.

Coefficient γ is the "valuation factor" introduced by Allcott and Wozny (2014). Consumers should be indifferent between spending a dollar on the up-front car purchase price and a dollar in present discounted value on total gasoline consumed (Allcott and Greenstone, 2012). Therefore, a value of $\gamma = 1$ implies that the consumers weigh the gas costs against the up-front car purchase prices in a fully rational way. A value of $\gamma < 1$ suggests that the consumers undervalue expected future gas costs at the time they initially purchase their cars. An overvaluation is indicated if $\gamma > 1$.

I define the present discounted value of a car's lifetime gas cost as an expectation over the gas price,

$$G_{ivt} = E\ [\ \sum_{s=1}^{S} \frac{1}{(1+r)^s} \frac{m_{it}}{mpg_v}\ g_{it}\ ], \tag{1.2}$$

in which g_{it} is the individual-specific gas price, m_{it} is the expected annual vehicle miles traveled of consumer i in market t, and mpg_v is the car-specific fuel economy rating. r is the interest rate at which consumers discount future gas costs, and S is the length of time over which consumers value gas cost savings generated from improved vehicle fuel efficiency.

Although I am using a static setup, a forward-looking new car buyer will divide this G_{ivt} into two parts: the present discounted value of gas costs during her holding period, and the present discounted value of gas costs over the remainder of the car's life after it is resold (Allcott and Wozny, 2014). In this model, car owner i knows the payout of selling this car to its next owner will be the present discounted value of the resale price plus that present discounted value of the remaining gas costs.

According to equation 1.2, G_{ivt} is heterogeneous if the individual-specific expected annual mileage is applied. I employ the market-specific empirical distribution of annual vehicle miles traveled D_t and draw m_{it} from corresponding D_t. Incorporating vehicle use heterogeneity in this way helps avoid the endogenous sorting bias identified by Bento et al. (2012) because consumers who drive more are likely to have a higher valuation of gas cost savings arising from increased vehicle fuel efficiency.

There could be multiple sources of consumer heterogeneity when modeling the present discounted value of the expected gas cost, such as a discount rate that varies across consumers and across time.[6] In this paper, I concentrate on heterogeneous annual mileage, so I assume the same interest rate r and time horizon S for all consumers.[7] By imposing such a simplifying assumption, I express the present discounted value of expected lifetime gas cost of car v as

$$G_{ivt} = \rho \, \frac{m_{it}}{mpg_v} \, E[g_{it}], \tag{1.3}$$

in which $\rho \equiv \sum_{s=1}^{S} (1+r)^{-s}$ is a common "capitalization factor" that measures how consumers trade off the up-front car purchase price against the expected annual gas cost over S years as in Grigolon et al. (2018). A fully myopic consumer will assign

[6] Multiple factors may affect consumer optimization in this context: a present bias, a systematically biased belief about the relative energy costs of products with different energy efficiency levels, or the inattention. DellaVigna (2009) provides a review of both the psychology and economics literature relevant here.

[7] As both low discount rate and long vehicle lifetime tend to result in consumer undervaluation, I employ different assumptions on these parameters so a list of alternative estimates of the valuation factor γ can be presented and analyzed.

no weight to annual gas cost, in which case the present discounted value of expected lifetime gas cost in equation 1.3 would be zero.

1.3.2 Individual Choice Probability and Market Share

Using equation 1.1 and equation 1.3, I rewrite the conditional indirect utility of consumer i for car v in market t as follows:

$$u_{ivt} = \alpha_i \ x_v + \beta_i \ p_v + \beta_i \gamma \rho \ \frac{m_{it}}{mpg_v} \ E[g_{it}] + \zeta_{vt} + \epsilon_{ivt}. \tag{1.4}$$

I use available data on x_v, p_v, m_{it}, mpg_v and $E[g_{it}]$ to estimate α_i, β_i, and the coefficient of expected annual gas cost $\beta_i\gamma\rho$. After that, I divide $\beta_i\gamma\rho$ by the price sensitivity β_i to obtain the product of the valuation factor and the capitalization factor $\gamma\rho$.

Since I cannot separately identify the valuation factor γ and the capitalization factor ρ, I take two steps to compute γ as in Grigolon et al. (2018). First, I apply assumptions on interest rate r and time horizon S to calculate ρ. Next, I divide the product $\gamma\rho$ by the capitalization factor ρ to reveal γ. Given the individual-specific annual mileage m_{it} and the fuel economy rating of car v, γ depends on the interest rate, the relevant time horizon, and the gas price expectation.

I let each new car buyer choose to purchase the alternative that delivers the highest level of utility relative to all other options in her choice set. I assume random coefficients α_i, β_i, and $\beta_i\gamma\rho$ come from a distribution $F(\theta)$ where θ includes all parameters of this distribution. Based on equation 1.4, the individual choice probability of consumer i for car v in market t is

$$s_{ivt}(\zeta_{vt}; \alpha_i, \beta_i, \beta_i\gamma\rho) = \frac{exp(\alpha_i \ x_v + \beta_i \ p_v + \beta_i\gamma\rho \ \frac{m_{it}}{mpg_v} \ E[g_{it}] + \zeta_{vt})}{1 + \sum_{v'=1}^{J} exp(\alpha_i \ x_{v'} + \beta_i \ p_{v'} + \beta_i\gamma\rho \ \frac{m_{it}}{mpg_{v'}} \ E[g_{it}] + \zeta_{v't})}. \tag{1.5}$$

Then the predicted market share of car v in market t is

$$S_{vt}(\zeta_{vt};\theta) = \int_{(\alpha_i,\beta_i,\beta_i\gamma\rho)} s_{ivt}(\zeta_{vt};\alpha_i,\beta_i,\beta_i\gamma\rho)\ dF(\theta). \tag{1.6}$$

1.3.3 Heterogeneous Consumer Response to Alternative Policies

Following Grigolon et al. (2018), I use the individual choice probability described in equation 1.5 to show that properly accounting for vehicle use heterogeneity is vital when investigating the relative effectiveness of the fuel tax and the product tax in reducing gasoline consumption. I let τ^g denote a new excise tax on retail gasoline sales which generates the change in fuel tax. Alternatively, I place a product tax on a car's per-mile fuel consumption (in gallons/mile). Since the per-mile fuel consumption is the inverse of a car's fuel economy rating (in miles/gallon, i.e., MPG), I write this product tax as $\frac{\tau^{mpg}}{mpg_v}$. Given these notations, the expected post-tax gas price is $E[g_{it}] + \tau^g$ per gallon and the post-tax vehicle purchase price becomes $p_v + \frac{\tau^{mpg}}{mpg_v}$.

Plugging new prices into equation 1.5, the probability of consumer i choosing model v in market t becomes

$$s_{ivt} = \frac{exp(\alpha_i\ x_v + \beta_i\ (p_v + \frac{\tau^{mpg}}{mpg_v}) + \beta_i\gamma\rho\ \frac{m_{it}}{mpg_v}\ (E[g_{it}] + \tau^g)) + \zeta_{vt})}{1 + \sum_{v'=1}^{J} exp(\alpha_i\ x_{v'} + \beta_i\ (p_{v'} + \frac{\tau^{mpg}}{mpg_{v'}}) + \beta_i\gamma\rho\ \frac{m_{it}}{mpg_{v'}}\ (E[g_{it}] + \tau^g)) + \zeta_{v't})}.$$

To inspect the consumer response to the implementation of a new gasoline tax and a product tax, I differentiate the individual choice probability with respect to both τ^g and τ^{mpg}, as explained in the Appendix. Consumer i's response given the change in gasoline tax is

$$\frac{\partial s_{ivt}}{\partial \tau^g} = \beta_i\gamma\rho m_{it}\ s_{ivt}\ (\frac{1}{mpg_v} - \sum_{v'=1}^{J} \frac{s_{iv't}}{mpg_{v'}}). \tag{1.7}$$

Similarly, consumer i responds to the product tax by adjusting her choice probability such that

$$\frac{\partial s_{ivt}}{\partial \tau^{mpg}} = \beta_i \; s_{ivt} \; (\frac{1}{mpg_v} - \sum_{v'=1}^{J} \frac{s_{iv't}}{mpg_{v'}}). \tag{1.8}$$

Since the price sensitivity β_i is negative, the common part in the parentheses of equations 1.7 and 1.8 implies that, when implementing either tax, the probability of consumer i choosing car v will decrease if the per-mile gasoline consumption of car v is higher than the share-weighted average per-mile gasoline consumption of all new car models in the market. This result fits intuition because the cost of purchasing and operating a car with high fuel economy rating will be lower than that cost for a less fuel-efficient car model after applying either τ^g or τ^{mpg}, all else equal.

The difference between consumers' responses to alternative policies lies in the term outside of the parentheses. The fuel tax influences the individual choice probability differently because the combination $\beta_i \gamma \rho m_{it}$ relates a consumer's response to her valuation of the expected gas cost savings. Moreover, this effect varies when there is heterogeneity in vehicle use intensity across new car owners.

In a set of policy simulations, I show that high-mileage drivers are more responsive to the new fuel tax than to the product tax despite the inelastic demand for vehicle miles traveled while low-mileage drivers are more responsive to the product tax. This heterogeneous consumer response provides the fuel tax with a clear advantage over the product tax in generating fuel savings because high-mileage drivers take up a more substantial proportion of the total gasoline consumption. More importantly, models not accounting for vehicle use heterogeneity understate the fuel saving effects of both policies and misleads us about relative effectiveness when comparing different policies.

In addition, after I relax the assumption of perfectly inelastic demand for driving with respect to changes in gasoline prices in policy simulations, the fuel tax proves to be a much more effective policy tool in reducing gasoline consumption of new car buyers relative to the product tax because the former influences consumer responses on both extensive and intensive margins.

1.4 Data

With the model and its assumptions in mind, this section presents the sample I construct using several data sources. The primary data used in this paper is the Massachusetts Vehicle Census (Metropolitan Area Planning Council, 2016). It is based on the Automated License and Registration System and a separate database containing records of vehicle inspections, both of which are administrative data sets maintained by the Massachusetts Registry of Motor Vehicles (Reardon et al., 2016).

1.4.1 Vehicle Choices

In Massachusetts, passenger vehicle registration renewal is valid for two years while vehicles are required to be inspected annually and within seven days of sale. When constructing the Massachusetts Vehicle Census (MAVC), registration records are split where a vehicle inspection record, which delivers the mileage reading, begins or ends. So each record in the MAVC covers a defined period when the specified vehicle had a unique combination of owner, garaging address, and average daily mileage (Reardon et al., 2016).

Vehicle manufacturer, model, fuel type, fuel economy rating, curb weight, and the manufacturer suggested retail price (MSRP) of each vehicle are included in the MAVC data. Vehicle identification number (VIN) and ZIP Code of garaging address are also available from the MAVC researcher files.

I apply a highly disaggregated definition of the vehicle model to capture the variation in fuel efficiency and engine performance as much as possible. Each vehicle recorded in the MAVC is a manufacturer/model/model year (MY) combination, e.g., “Volkswagen Jetta 2011”. I use VINs and the VIN decoder provided by the National Highway Traffic Safety Administration (NHTSA) to retrieve the trim level information of each vehicle.[8] Also, I use the trim level information to collect extra vehicle attributes such as vehicle body type, vehicle passenger volume, and interior cargo volume from Cars.com and Ward's Automotive Yearbook. The unit of observation in my sample is at the very detailed level, e.g., "Volkswagen Jetta 2011, 2.5 liter, 170 hp, 3,045 lbs, 27 MPG, 94.1 ft^3 passenger volume, and 15.5 ft^3 cargo volume".

1.4.2 New Car Markets

In Massachusetts, thirteen Metropolitan Planning Organization (MPO) and Transportation Planning Organization (TPO) regions cover all municipalities of the state. An MPO/TPO is a federally required regional transportation policy-making organization made of representatives from local government, regional transit operators, and state transportation agencies. Each MPO/TPO creates a fair and impartial setting for effective regional decision making in the metropolitan area to coordinate and manage transportation projects and programs that carried out in 351 towns and cities across the state. Following the Massachusetts Travel Survey published by the Massachusetts Department of Transportation in 2012, I employ MPO and TPO regions to define geographic new car markets in Massachusetts.

I choose my sample period to be the model year overlapping months during which both new MY 2011 and 2012 cars are for sale. Based on disaggregated new vehicle transaction data, Copeland et al. (2011) point out that for about half the calendar year, automakers simultaneously sell two vintages of the same model. Table 1.1 shows that about 95 percent of the statewide new passenger vehicle registrations during the sample period is for MY 2011 and 2012.

The first MY 2012 car observed from the MAVC is in March 2011, but the 2012 vintage does not show up on a large scale until July 2011. Although upcoming model

[8]Manufacturers use trim levels to identify a vehicle's level of equipment or special features. For models that use several trim choices, automakers usually offer three or four versions. For example, the gasoline-powered 2011 Volkswagen Jetta comes in three versions: S, SE, and SEL. The Jetta S is the base model, which includes the fewest features and has the lowest price of the three. The SE is in the middle of the range in both price and equipment, and the SEL is the most luxurious and feature-rich version.

year debuting vehicles can be spotted in as early as March of the current calendar year, August and September are generally when automakers transition to the new model year (Copeland et al., 2011). According to the availability of MY 2012, I use new car registration records of MY 2011 and MY 2012 from ten MPO/TPO regions in Massachusetts between July 2011 and December 2011 to construct the monthly new car market.[9]

This batch of observations has the longest inspection history within the MAVC time frame for me to trace mileage records. Also, as required by the tightened fuel economy standards of passenger cars under the Obama administration, the sales-weighted fleet average fuel economy rating of passenger cars increases from 30.4 MPG for MY 2011 to 33.3 MPG for MY 2012 (EPA and NHTSA, 2010). Such changes in vehicle fuel efficiency provide extra identification source for me to pin down consumer valuation of expected fuel cost savings.

1.4.3 Prices

New Car Prices. I use MSRPs for new car purchase prices because transaction prices are not available. Market dummies are included in the mean utility term of each car model to account for manufacturer and dealer incentives offered to the car buyers and other market fixed effects. Using monthly new vehicle transaction prices, Copeland et al. (2011) document that the average retail price of a new vehicle model declines at an annual rate of nine percent over its model year. To bring this feature of new vehicle markets to my sample, I apply a nine percent discount to MSRPs of all MY 2011 cars to create this "new vintage premium" for MY 2012 cars.

Gasoline Prices I employ the average price of midgrade motor fuel in Massachusetts from the U.S. Energy Information Administration (EIA) to form the consumer expectation of future gasoline prices.[10] The EIA uses a regional classifica-

[9]I drop three MPO regions when building the sample because there are very few observations in each of them. See the Appendix for a detailed data cleaning procedure.

[10]I also apply prices of the regular grade gasoline for robustness check. The results are nearly identical.

tion that divides the U.S. into seven Petroleum Administration for Defense Districts (PADD). It provides gas price information for these seven regions, and also for ten states and nine cities separately. Figure 1.1 shows that the monthly average price of the midgrade gasoline in the New England area (PADD1A), the statewide average price of Massachusetts, and the citywide average price of Boston share the similar pattern. The latter two overlap because these two average prices are very close to each other. The prices that I use for computing the average price of midgrade gasoline lie between two vertical dotted lines.

The expected gasoline price of consumers plays a vital role in my model. Based on the Michigan Survey of Consumers (MSC), Anderson et al. (2011) conclude that households in the MSC typically form expectations about the inflation-adjusted price of gasoline using a simple no-change model. Specifically, households surveyed in MSC expect the nominal gasoline price to grow at the same rate as inflation, which is equivalent to expecting the real price of gasoline in the future to be the same as the current price of gasoline. Although the inflation expectations data in the MSC is limited to selected horizons, this no-change model is valid when advanced methods of inflation forecasting are employed (Baumeister and Kilian, 2016). I adopt this simple rule of thumb in my analysis. Therefore, the consumer gas price expectation in my model is the average of Massachusetts midgrade gasoline price over six months from July 2011 to December 2011.[11]

1.4.4 Empirical Distribution of Annual Vehicle Miles Traveled

I construct the ownership and inspection history for observations in my sample from the vehicle registration and inspection records contained in the MAVC. Each valid vehicle inspection record in the MAVC reports the number of days between two inspections and the average daily miles traveled during that period. I weight the

[11] In addition to applying this no-change model, I evaluate several alternative behavioral models of consumers' gas price expectations following Kilian and Sims (2006). These alternative models of consumers' expected gas prices give similar results to the no-change model.

daily miles using the length of the time span between two inspections, and apply this weighted average daily mileage to calculate the expected annual vehicle miles traveled.

From the new car market sample, I select all registered new cars with complete inspection records over three consecutive years and apply the weighting procedure described above to their mileage records to create the mileage sample. Table 1.2 presents the numbers of observations in both the demand sample and the mileage sample by markets. I also report summary statistics of daily vehicle miles traveled by markets in table 1.2. Overall, the average number of the days between two inspections is 379, and the standard deviation is 32.

1.5 Empirical Implementation

I employ a sequence of T markets to estimate the taste parameters in a system of market shares using the Generalized Method of Moments (GMM). The data I am using has three components: car attributes including price, market shares, and empirical distributions of the annual mileage.

1.5.1 Specification of the Taste Parameters

To achieve flexible substitution patterns for new car buyers, I allow the marginal utilities of some observed car attributes and the expected gas cost to vary at the individual level. The conditional indirect utility of consumer i for car v in market t becomes

$$u_{ivt} = \bar{\alpha}\ x_v + (\Sigma^K \nu_i)'\ x_v^K + \beta\ p_v + \beta_i \gamma \rho\ m_{it}\ \frac{g}{mpg_v} + \zeta_{vt} + \epsilon_{ivt} \tag{1.9}$$

in which $\bar{\alpha}$ is a vector of mean valuations for all observed car attributes, ν_i is a $K \times 1$ vector of unobserved (to the econometrician) idiosyncratic tastes for K observed car attributes x_v^K, and Σ^K is a matrix with parameters σ^k on the diagonal.[12] I apply K independent $\chi^2(3)$ distributions truncated at 95 percent for ν_i following Petrin (2002).[13] Parameters σ^k capture the heterogeneity in unobserved tastes ν_i for x_v^K in the population.

p_v is the MSRP of new car model v. I apply the average Massachusetts midgrade gasoline price over the sample period to all consumers. This approximation implies $E[g_{it}] = g$ for every new car buyer in each MPO market. I specify g/mpg_v as the per-mile gas cost of driving (in dollars per mile).

The product $\beta_i \gamma \rho$ is estimated as one random coefficient with draws m_{it} from corresponding market-specific mileage distribution D_t. These empirical mileage distributions work as distributions of demographic characteristics obtained from representative data sets such as the Current Population Survey in providing extra information and therefore ensuring the identification of $\beta_i \gamma \rho$.

The term ζ_{vt} represents unobserved car attributes. It does not have a random coefficient. The inclusion of this term allows the model to rationalize patterns of market shares observed from the data. I define this term of unobserved car attributes as

$$\zeta_{vt} = \zeta_t + \tilde{\zeta}_{vt} \tag{1.10}$$

in which ζ_t captures the market fixed effects, and $\tilde{\zeta}_{vt}$ are random effects accounting for any remainder of unobserved product attributes that vary across different car models and markets.

[12]For simplicity, I restrict the covariances of all random coefficients to be zero as in Nevo (2000).

[13]This distributional assumption implies that the heterogeneity of unobserved consumer tastes in the population is skewed in the positive direction. When more information becomes available, I can construct full demographic-dependent taste terms. I also apply the standard normal distribution for ν_i for a robustness check. The results obtained from this alternative distributional assumption are similar.

1.5.2 Identification and the GMM Estimator

To incorporate vehicle use heterogeneity, I apply the discrete choice model to define the individual choice probability and hence the market share of a particular new car model as a function of the car attributes (including price) and the new car buyer's expected gas cost. In my model, a consumer knows what tasks her newly purchased passenger car will perform before she buys it. Therefore, the vehicle use pattern of a car bought by consumer i is consistent over her holding period. The variation in the up-front purchase price and the expected gas cost across different car models allows me to identify how much new car buyers value the expected fuel cost savings arising from improved vehicle fuel efficiency when they initially purchase their cars.

To pin down the parameters governing the substitution patterns of the new car buyers, I need variations in attributes across new car models and expected gas cost across consumers. As summarized in table 1.2, market-specific empirical mileage distributions help capture the variation in vehicle use pattern among consumers in different markets. Since the gasoline price is the same for all new car buyers, the variation in the expected car lifetime gas cost across drivers in my sample comes from the difference in the model-specific vehicle fuel efficiency and the individual-specific expected annual vehicle miles traveled rather than from changes in gasoline prices. I present the summary statistics for attributes of new car models purchased by consumers in table 1.3. Attributes that vary at the vehicle trim level such as car purchase price, performance ratio (i.e., horsepower/curb weight), fuel economy rating, and engine size have higher dispersion around the mean than attributes changing at the nameplate-model level like the interior passenger volume.

Following the seminal work Berry (1994), I define the product-specific linear mean utility component of u_{ivt} as

$$\delta(S_{vt}, \theta) \equiv \bar{\alpha}\ x_v + \beta\ p_v + \zeta_t + \tilde{\zeta}_{vt} \tag{1.11}$$

in which S_{vt} is the market share of model v in market t. Given the distributional assumption on consumer taste for car attributes and the expected gas cost, Berry (1994) demonstrates that there is a unique $\delta_{vt}(\theta)$ that solves $S_{vt}^{data} - S_{vt}(\delta, \theta) = 0$ for each θ, in which S_{vt}^{data} is the observed market share.

I address price endogeneity by using price shifters z_{vt} in market t as instrumental variables. In particular, for car v, the squares of its own attributes, the sum of each attribute of car models made by the same manufacturer, and that of car models made by competing manufacturers are employed as instrumental variables. These instruments are introduced by Pakes (1994), and have been used by Berry et al. (1995) and subsequent work.[14]

By applying this group of instruments, I assume that the unobserved car attributes $\tilde{\zeta}_{vt}$, although correlated with the car purchase price, are mean independent of those observed nonprice car attributes and the gas price expectation. I rule out any endogeneity in observed nonprice attributes following previous literature, which appears reasonable because manufacturers cannot quickly redesign their products over such a short span of time as I am considering in this paper. The cost of maintaining this assumption is the possibility of inconsistent estimators if it turns out to be inappropriate (Greene, 2011).

I compute $\delta(S_{vt}, \theta)$ numerically using the contraction mapping procedure developed by Berry et al. (1995). Given this mean utility term and the assumption $E(\tilde{\zeta}_{vt}(\theta)|z_{vt}) = 0$, equation 1.11 becomes a standard linear regression model. The

[14]This set of price shifters are the standard instruments used in random coefficients logit demand applications (Gandhi and Houde, 2016). They have been proved very effective in the study of many industries such as automobiles, computers, and pharmaceutical drugs (Nevo, 2000). For automobiles, previous papers that apply price shifters include Petrin (2002), Berry et al. (2004), and Train and Winston (2007).

primary set of moments that I use for the GMM estimator is $E(\tilde{\zeta}_{vt}(\theta) z_{vt}) = 0$.[15] The GMM estimator $\hat{\theta}_{gmm}$ is the solution to this following criterion function

$$\hat{\theta}_{gmm} = \underset{\theta}{\operatorname{argmin}}\, \tilde{\zeta}(\theta)' Z W^{-1} Z' \tilde{\zeta}(\theta) \tag{1.12}$$

in which Z is the instrument matrix and I set the weighting matrix W to be $Z'Z$ following Nevo (2000). I take into account both sampling error and simulation error to estimate the standard errors of the parameter estimates following Hansen (1982). I describe the detailed estimation procedure and the numerical considerations in the Appendix.

1.6 Results

Correctly accounting for consumer heterogeneity has been recognized as an essential modeling feature when estimating the demand for differentiated products in general (Ackerberg et al., 2007). When examining the consumer valuation of expected gas cost savings, I show that properly accounting for heterogeneity in vehicle use intensity helps address the endogenous sorting of low-mileage drivers into less fuel-efficient car models while considering heterogeneity in tastes for observed car attributes alone cannot achieve this correction. To do so, I compare parameter estimates and corresponding valuation factors from four different demand models.

The first model (IV Logit Mean Miles) is an instrumental logit model combining price shifters as instrumental variables to address the price endogeneity. In this simple logit model, the expected annual vehicle miles traveled equals the mean of the observed mileage distribution for consumers in each market, and there is no heterogeneity in consumer taste for observed car attributes.

[15] According to the Metropolitan Area Planning Council of Massachusetts (MAPC), the MAVC database is currently undergoing an update. A more extended time series of the data (through June/2017) with more attributes (e.g., owner birth year) can be expected soon. In that case, I will augment micro moments to the current estimation procedure for additional sources of identification following Petrin (2002).

To investigate changes in the demand estimation and the valuation factor resulted from including different types of heterogeneity, I develop the second model (IV RCLogit Mean Miles) and the third model (IV Logit EPD Miles) to account for taste heterogeneity and vehicle use heterogeneity accordingly.[16] The IV RCLogit Mean Miles model adapts random coefficients for consumer preference over some of the observed car attributes as laid on in the previous section while keeps using mean mileage. The IV Logit EPD Miles model builds on previous IV Logit Mean Miles model. It introduces mileage heterogeneity by incorporating random draws from the market-specific empirical mileage distributions. Therefore, heterogeneous annual vehicle miles traveled are employed to calculate the expected future gas cost while the valuation of car attributes remains homogeneous.

Finally, the last demand model (IV RCLogit EPD Miles) applies random coefficients to the expected gas cost and also the observed car attributes. It allows for heterogeneity in both vehicle miles traveled and consumer valuation of observed car attributes.

1.6.1 Parameter Estimates

Table 1.4 reports parameter estimates from four demand models. Results from the IV Logit Mean Miles model (column 1 in table 1.4) are intuitive. New car buyers have strong disutility from high car purchase prices, and they penalize high expected car lifetime fuel costs. While omitted in table 1.4, parameter estimates on market dummies have consistent patterns across all four demand models. They suggest that residents living in any MPO regions but Boston are more likely to buy gasoline powered cars relative to consumers living in the Greater Boston area.

The average own-price elasticity of new car demand across ten MPO markets is -1.31.[17] This number indicates a less elastic demand in comparison with elasticities

[16] RCLogit stands for Random Coefficients Logit Model. EPD is short for Empirical Distribution.

[17] Elasticity calculated from a logit model without using the instrumental variables is about -0.04. When applying the price shifters as instruments, the F-statistic of the first-stage mean utility re-

estimated in previous literature where the average elasticity across vehicle segments is in the neighborhood of -2 to -5.[18] As conventional wisdom would posit, price elasticity diminishes with income. Inelastic demand for the new cars in my sample is reasonable because consumers in Massachusetts enjoy higher household incomes relative to the rest of the country. Additionally, the consumer demand for passenger cars is less elastic than that for SUVs and pickup trucks in general (Copeland et al., 2011).

Both the IV RCLogit Mean Miles model (column 2 in table 1.4) and the IV Logit EPD Miles model (column 3 in table 1.4) deliver the price sensitivity parameter close to that from the IV Logit Mean Miles model. When allowing for heterogeneity in annual mileage and hence in the valuation of expected gas cost in the IV Logit EPD Miles model, the impact of expected gas cost on utility entirely varies with the individual-specific annual mileage, so the mean valuation is omitted. The expected gas cost again have a negative and significant effect on consumer utility. According to previous literature, horsepower and size are top essential attributes for U.S. consumers to consider when buying new vehicles. The estimated mean valuations in column 3 indicate that new car buyers in Massachusetts also favor cars with rapid acceleration and large interior passenger volume. Moreover, they have a strong preference for European cars compared to those made by American manufacturers.

As I apply heterogeneity to both annual vehicle miles traveled and consumer valuation of car attributes in the full random coefficients logit model (column 4 in table 1.4), the effects of the car purchase price and gas cost on utility are consistently negative and significant. However, when comparing IV RCLogit models to IV Logit models, it seems that more demographic information on the population is required to

gression is 49. According to the bias method introduced by Stock and Yogo (2002), I reject weak instruments at the 95% confidence level.

[18] Goldberg (1995) reports residual demand elasticities of demand for specific vehicles that are in the -2 to -4 range. Berry et al. (1995) report elasticity estimates for 13 vehicle models with the average of -5. Using recent data, Bento et al. (2009) report -1.9 for all new passenger vehicles, and Whitefoot et al. (2011) report -1.97.

identify consumer preference over car performance and passenger volume if random coefficients are allowed for these attributes.

1.6.2 Consumer Valuation of Expected Gas Cost Savings

In this section, I calculate consumer valuation of expected gas cost savings arising from improved vehicle fuel efficiency using parameter estimates obtained from demand models that employ different mileage information. I show that vehicle use heterogeneity must be included to reveal the true valuation factor γ if bias is caused by the endogenous sorting of low-mileage drivers into less fuel-efficient car models.

I divide the gas cost coefficient $\beta_i\gamma\rho$ by the price coefficient β to obtain the product of the valuation factor and the capitalization factor $\gamma\rho$. Then I apply assumptions on the discount rate and the time horizon to compute the capitalization factor ρ. Finally, I take the ratio of $\gamma\rho$ over ρ to reveal the valuation factor γ. My calculation is closer to the full-valuation results reported by Busse et al. (2013) and Grigolon et al. (2018) than to the partial-valuation results obtained by Allcott and Wozny (2014) and Leard et al. (2017).

A reasonable estimation of the valuation factor γ relies on choices carefully made for the interest rate r and the time horizon S. Allcott and Wozny (2014) calculate the average discount rate weighted over used vehicle buyers using different payment methods (i.e., financed, leased, and cash) and set the discount rate as 6%. I adopt this value of the discount rate for analysis.

The vehicle sustainability and travel mileage schedules published by the NHTSA suggests that the average maximum vehicle age of passenger cars is 25 in the U.S. (Lu, 2006). Using the 2009 NHTS data, Leard et al. (2017) updates the maximum lifespan for cars to 35 years as better technology and overall vehicle quality improvements have been driving up the average vehicle age over time. I apply both numbers of the car lifetime in my calculation.[19]

The first row of column 1 in table 1.5 shows that the product $\gamma\rho$ calculated from the IV Logit Mean Miles model is noticeably smaller than those obtained from models accounting for heterogeneity among consumers. Consequently, in the second and the third row of table 1.5, the valuation factors computed from the IV Logit Mean Miles model predict a noticeable consumer undervaluation of expected gas cost savings arising from improved vehicle fuel efficiency. This corresponds to the conclusion reached by Bento et al. (2012) through a simulation study: when the undervaluation of energy costs is not present in the true data generating process, the simple logit model could erroneously suggest significant undervaluation while the random coefficients logit model recovers the actual value.

Columns 2 and 3 in table 1.5 present similar results. When choosing time horizon as 35 years, the models accounting for heterogeneity in either consumer preference over car attributes or expected annual mileage indicate a moderate consumer undervaluation of expected gas cost savings. After including heterogeneity in both dimensions, the valuation factor results presented in column 4 suggest that consumers fully value the benefits of improved vehicle fuel efficiency when they initially purchase their cars (i.e., $\gamma = 1$).

I also calculate the implicit discount rate for new car buyers following Hausman (1979). This implicit discount rate is the interest rate at which consumers discount the future, assuming that they value the expected gas cost savings to the full extent over a given value of the car lifetime. In the last two rows of table 1.5, the implicit discount rates calculated from the full model employing random coefficients for both car attributes and expected annual mileage are within a reasonable distance from 6%, which is the discount rate calculated by Allcott and Wozny (2014).

[19] A relatively long average vehicle lifetime tends to conclude consumer undervaluation. The definition of the capitalization factor is $\rho \equiv \sum_{s=1}^{S} (1+r)^{-s} = \frac{1}{r}[1-(1+r)^{-s}]$. When the product $\gamma\rho$ is fixed, a lower r or a larger S (i.e., higher ρ) leads to a smaller γ.

Although the gasoline price was trending down during the sample period, it was still at a high point relative to gasoline prices during the previous years. In this case, consumers seem to behave rationally by assigning a near-full valuation to expected gas cost savings. Hassett and Metcalf (1993) argue that high energy expenditure increases the return to an energy saving investment because the return to the investment is the energy cost avoided. Similarly, Busse et al. (2016) suggest that, if consumers experience an adjustment cost to changing their vehicle choices or vehicle use patterns, they may not respond to changes in gasoline prices until the price crosses a threshold at which it becomes worthwhile to make the necessary switch. This might, therefore, serve as the motivation to implement the gasoline tax properly or apply a retail sales tax to gasoline purchases.[20]

1.7 Policy Simulations

At the root of designing effective policies in producing fuel savings and emissions reductions in passenger transportation is the influence of policies on consumer behavior. In this section, I carry out policy simulations to show that accounting for vehicle use heterogeneity is critical when evaluating such policies because it more accurately captures the heterogeneous consumer response to policy changes, which provides the fuel tax with a clear advantage over the product tax in reducing gasoline consumption. To do so, I apply random coefficients demand models employing different mileage information to compare the amounts of fuel savings resulting from implementing alternative policies.

1.7.1 Perfectly Inelastic Demand for Driving

Policy simulations conducted in this section investigate the effects of a fuel tax and a product tax on consumers' vehicle choice decisions but not on how many miles

[20]States currently imposing a retail sales tax for gasoline fuel purchases are California, Florida, Georgia, Hawaii, Illinois, Indiana, Michigan, New York, West Virginia.

to drive. In particular, while the fuel tax may directly affect consumers' demand for vehicle miles traveled because it changes the cost of driving, I fix the expected annual mileage before and after applying new taxes for every simulated new car buyer. This implementation is consistent with the literature where most researchers find relatively inelastic consumer demand for the vehicle miles traveled with respect to changes in gasoline prices (Gillingham, 2014; Davis and Kilian, 2011). Moreover, according to recent studies on the "rebound effect" in the context of transportation equipment, an improvement of the vehicle fuel efficiency hardly influences the vehicle use pattern of consumers (Gillingham et al., 2016).

In column 1 of table 1.6, I report the current average annual gasoline consumption of new car buyers in the Boston area that calculated from observed mileage records and market shares of new cars. Columns 2 and 3 of table 1.6 present the amounts of expected average annual gasoline consumption per driver when I in turn apply a counterfactual increase of the fuel tax by \$0.25/gallon to the average midgrade gasoline price in Massachusetts, which is \$3.69/gallon over the sample period, and a revenue equivalent product tax on the per-mile gasoline consumption of a car model.[21]

The first row in Panel A of table 1.6 shows that, in a model accounting for heterogeneity in consumer preference over car attributes but not vehicle usage, a 25 cents increase in the fuel tax has an almost identical effect as a revenue equivalent product tax in reducing gasoline consumption of new car buyers. The expected gasoline consumption decreases to about 387 gallons per year per driver in both scenarios. However, the second row in Panel A suggests that in a model properly accounting for vehicle use heterogeneity, a fuel tax reduces the expected average annual gasoline consumption by 1.37 percent while a revenue equivalent product tax reduces the consumption by only 0.90 percent.

To investigate the mechanism that generates this relative effectiveness result, I examine the impacts of alternative policies on different groups of drives. I assign simulated new car buyers who land in the top half of the mileage distribution to the

[21]Refer to the Appendix for the calculation of a revenue equivalent product tax.

"High Miles" group and the rest to the "Low Miles" group. When considering only low-mileage drivers (the first row in Panel B of table 1.6), the counterfactual fuel tax reduces about the same annual gasoline consumption on average as the product tax does. However, the fuel tax has a more substantial fuel-saving effect on high-mileage drivers. When applying an extra fuel tax \$0.25/gallon, high-mileage drivers reduce their average annual gasoline consumption from about 508 gallons to roughly 499 gallons by switching to more fuel-efficient cars. Although a revenue equivalent product tax also leads to trimming in average annual gasoline consumption of high-mileage drivers, as suggested by numbers in the second row in Panel B of table 1.6, the cutback is smaller.

This set of comparisons yields important insights. When consumers fully value the gas cost savings arising from improved vehicle fuel efficiency (i.e., $\gamma = 1.01$), a fuel tax and a product tax incentivize different groups of consumers to substitute toward more fuel-efficient car models and therefore generate differentiated fuel saving results. When breaking down the impacts of alternative policies by mileage group, I show that high-mileage drivers are more responsive to policy changes relative to low-mileage drivers. More importantly, high-mileage drivers are more responsive to a change in fuel taxes than to a change in product taxes, while low-mileage drivers are more responsive to product taxes. In addition, I show that when the same set of policies is considered, a model not employing heterogeneous miles not only understates the fuel saving effects of both policies but also mislead us about the relative effectiveness of two policies in reducing gasoline consumption of new car buyers.

According to data provided by the U.S. Department of Energy, the national average annual fuel use per vehicle of the passenger car category is 480 gallons as of 2015, which is close to the average annual gasoline consumption of high-mileage drivers in my sample. Therefore, the aggregate fuel savings and emissions reductions generated from correctly implementing fuel taxes could be considerable at the national level even when the demand for driving is held constant.

1.7.2 Elastic Demand for Vehicle Miles Traveled

In the previous section, I assume a perfectly inelastic demand for driving to show that a fuel tax is more effective in reducing gasoline consumption as it incentivizes high-mileage drivers more to choose fuel-efficient cars. In this section, I allow consumers to respond to changes in fuel prices by adjusting how much to drive. Therefore, the second set of policy simulations carried out in this section also account for the effect of changes in fuel taxes on consumer response on the intensive margin.[22] In the meantime, I stick to the "weak rebound effect" result from recent literature, which indicates that consumers do not drive more after switching to more fuel-efficient vehicles, so new car buyers' decisions on how much to drive are held constant even after implementing the product tax.

I follow the same procedure conducted in the previous section but apply a new mean mileage number or a new mileage distribution accordingly after implementing the extra fuel tax. Numbers in the first row in Panel C of table 1.6 suggest that, even when not accounting for heterogeneous vehicle usage among consumers, an extra fuel tax of \$0.25/gallon reduces more gasoline consumption relative to a revenue equivalent product tax when drivers are allowed to switch vehicles and to adjust mileage in response to changes in fuel taxes.[23]

In a model accounting for heterogeneity in both consumer preference over car attributes and vehicle usage, the difference in effects on reducing gasoline consumption between a fuel tax and a revenue equivalent product tax is more dramatic. The new fuel tax reduces about ten more gallons per year per driver compared to the product tax. When looking into different driver groups in Panel D of table 1.6, both low-

[22]Using mileage information obtained from the MAVC vehicle registration and inspection records, I create a vehicle level panel data that includes about 6,000 new passenger cars over three years. Following Gillingham et al. (2015), I estimate the elasticity of driving with respect to gasoline prices and predict new miles after applying counterfactual fuel taxes. Refer to the Appendix for estimation details and results.

[23]In Panel C, the mean mileage in the IV Mean Miles model changes after applying a new fuel tax. Numbers of the percentage change in the average annual gasoline consumption for the product tax are slightly different for mean miles models in Panel A and Panel C because I anchor product taxes on fuel tax revenues.

mileage drivers and high-mileage drivers make pronounced responses to an increase in the fuel tax by reducing their annual gasoline consumption. This set of policy simulations indicates that when accounting for its effects on consumer responses on both extensive and intensive margins, a change in fuel taxes has a clear advantage over a change in product taxes in reducing gasoline consumption of new car buyers, even when the magnitude of tax increase is small.

1.8 Concluding Remarks

I have shown that accounting for vehicle use heterogeneity is critical when evaluating the effectiveness of a fuel tax and a product tax in generating fuel savings and emissions reductions by analyzing models employing distributional miles and models considering only mean mileage. I construct the ownership and inspection history for new passenger cars using vehicle registration and inspection records contained in the MAVC, which covers almost every single vehicle registered in Massachusetts between 2009 and 2014. This high-quality vehicle level microdata allows me to form and apply distributions of expected annual mileage in demand estimation and policy simulations. The observed annual vehicle miles traveled during years after the new car purchase provide a better approximation of the car's lifetime mileage and facilitate the decomposition of the impacts of alternative energy policies on new car buyers.

I find that it is crucial to account for heterogeneity in both consumer preference over car attributes and vehicle use intensity when estimating consumer demand for new passenger cars. I show that properly accounting for vehicle use heterogeneity helps address the endogenous sorting of low-mileage drivers into less fuel-efficient car models while considering heterogeneity in tastes for observed car attributes alone can not achieve this correction.

My demand estimation indicates that consumers fully value gas cost savings arising from improved vehicle fuel efficiency when they initially purchase their cars. My policy simulations suggest that a fuel tax is more effective in reducing gasoline consumption because high-mileage drivers are more responsive to a change in fuel taxes than to a change in product taxes, even as low-mileage drivers are more responsive to product taxes. By capturing such heterogeneous consumer responses to policies, I show that a moderate increase in fuel taxes is more effective than a revenue equivalent product tax in reducing the total gasoline consumption of new car buyers even when the demand for driving is held perfectly inelastic. Moreover, a model not accounting for vehicle use heterogeneity understates the fuel saving effects of both policies and misleads us about the relative effectiveness of two policies.

Figures And Tables

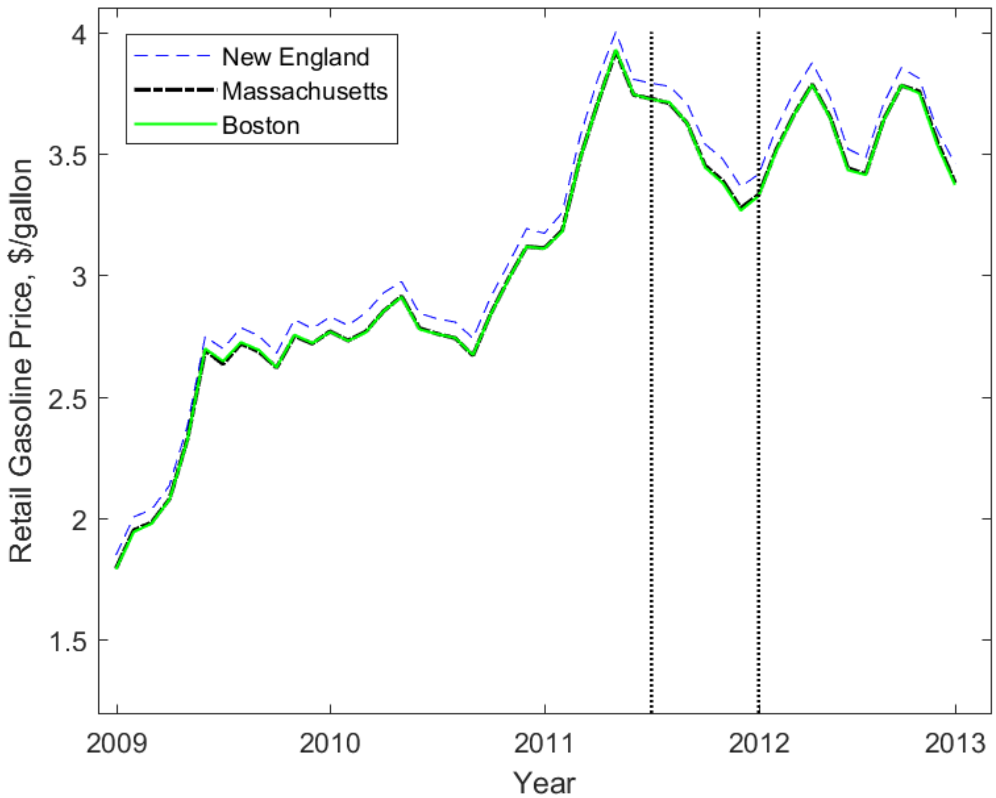

Figure 1.1. The monthly average price of midgrade gasoline in the New England area, the statewide average price of Massachusetts, and the citywide average price of Boston are plotted in this figure. The latter two prices are visually indistinguishable because these two are similar. Prices that used to construct the state average price for empirical analysis lie between two vertical dotted lines. All prices are normalized to 2011 dollars. Source: EIA Monthly Gasoline Prices Table.

Table 1.1.
New Passenger Vehicle Registration Records

Model Year	**Number of Obs.**	**Percentage**
prior to 2010	1,182	5.82
2011	11,759	57.91
2012	7,366	36.27

Notes: These are the numbers of new passenger vehicles registered between the date when the first MY 2012 model is observed (March/20/2011) and the date right before the first MY 2013 model is logged (February/06/2012) in the MAVC.

Table 1.2.
Demand Sample and Mileage Sample

MPO/TPO Market	(1) Num. of Towns/Cities	(2) Demand Sample Obs.	(3) Mileage Sample Obs.	(4) Daily Veh. Miles Traveled Mean	Median	Max.	S.D.
Berkshire MPO	32	180	114	34.55	31.36	90.80	17.77
Boston Region MPO	97	5,104	3,532	31.20	29.32	124.53	15.46
Cape Cod MPO	15	297	197	39.77	37.66	153.63	22.37
Central MA MPO	40	717	493	39.80	36.90	120.13	19.30
Merrimack Valley MPO	15	409	297	41.51	40.49	119.83	18.40
Montachusett MPO	22	281	197	44.24	39.24	113.01	21.98
Northern Middlesex MPO	9	402	277	36.94	34.59	111.64	17.90
Old Colony MPO	17	461	323	38.88	37.48	109.68	18.57
Pioneer Valley MPO	43	645	444	35.94	33.05	136.01	18.48
Southeastern MA MPO	27	633	440	40.97	37.36	130.00	21.69
Total	**351**	**9,215**	**6,377**	**34.75**	**31.93**	**153.63**	**17.93**

Notes: MPO represents the Metropolitan Planning Organization, TPO stands for Transportation Planning Organization. The number of towns and cities in column (1) shows how many municipalities are included in a corresponding MPO or TPO. Column (2) reports the number of observations that involved in demand estimation. Column (3) tells the number of observations that have three consecutive mileage records over the sample period. Mean Daily Mileage column tells the mean value of average daily miles traveled per vehicle.

Table 1.3.
Vehicle Attributes

Variable	**Mean**	**Std. Dev.**
Vehicle Price	23,318.50	7013.14
Fuel Economy (MPG)	28.16	5.70
Engine Horsepower (bhps)	171.66	51.63
Engine Size (liter)	2.28	0.62
Curb Weight(lbs)	3,153.43	393.94
Passenger Volume (ft^3)	98.19	8.39
Cargo Volume (ft^3)	17.78	7.05
Premium Features (0-1)	0.06	0.23
Number of Observations	**9,215**	

Notes: The "premium features" refers to new cars cost more than $36,000. All prices are normalized to 2011 dollars.

Table 1.4.
Parameter Estimates from Alternative Models

	(1) IV Logit Mean Miles	**(2) IV RCLogit Mean Miles**	**(3) IV Logit EPD Miles**	**(4) IV RCLogit EPD Miles**
		Average Utility		
Constant	-7.12***	-6.63*	-6.92***	-6.37*
	(0.43)	(3.51)	(0.44)	(3.45)
Car MSRP	-0.57***	-0.51**	-0.55***	-0.48**
	(0.14)	(0.22)	(0.13)	(0.21)
Gas Cost	-6.45***	-6.57***	-	-
	(1.68)	(1.58)	-	-
Horsepower/Weight	0.91***	-1.67	0.90***	-2.23
	(0.51)	(5.09)	(0.28)	(5.84)
Passenger Volume	4.50***	4.91	4.48***	4.80
	(0.58)	(6.83)	(0.56)	(11.32)
Premium Features	0.90**	0.72	0.84**	0.64
	(0.67)	(0.61)	(0.35)	(0.53)
European Automaker	0.41**	0.44*	0.43**	0.47**
	(0.19)	(0.25)	(0.21)	(0.24)
Japanese Automaker	0.01	0.02	0.02	0.03
	(0.10)	(0.13)	(0.11)	(0.14)
Korean Automaker	-0.01	0.05	0.01	0.06
	(0.13)	(0.21)	(0.12)	(0.20)
		Utility that Varies over Consumers Related to Mileage Distribution		
Gas Cost/Income	-	-	-7.00***	-7.10***
	-	-	(1.76)	(1.80)
		Utility that Varies over Consumers Following $\chi^2(3)$		
Horsepower/Weight	-	0.58	-	0.67
	-	(0.81)	-	(0.88)
Passenger Volume	-	-0.31	-	-0.24
	-	(5.60)	-	(7.40)

Notes: RCLogit stands for Random Coefficients Logit Model. EPD is short for Empirical Distribution. 1,000 Monte Carlo draws and quasi-Newton method are used for random coefficients models. Standard errors are in parentheses. Significance levels for the Z-test are as follows: ***1 percent, **5 percent, *10 percent.

Table 1.5.
Consumer Valuation of Expected Gas Cost Savings

	(1) IV Logit Mean Miles	**(2) IV RCLogit Mean Miles**	**(3) IV Logit EPD Miles**	**(4) IV RCLogit EPD Miles**
Gas Cost/Price				
$\gamma\rho$	11.23	12.95	12.65	14.69
Valuation Factor				
γ (r = 6%, S = 25)	0.88	1.01	0.99	1.15
γ (r = 6%, S = 35)	0.77	0.89	0.87	1.01
Implicit Discount Rate (%)				
r (γ = 1, S = 25)	7.42	5.86	6.11	4.59
r (γ = 1, S = 35)	8.37	7.00	7.22	5.89

Notes: RCLogit stands for Random Coefficients Logit Model. EPD is short for Empirical Distribution. The combination $\gamma\rho$ is computed using parameter estimates from the demand estimation. r is the discount rate. S is the maximum car lifetime.

Table 1.6.
Expected Average Annual Gasoline Consumption

	Expected Average Annual Gasoline Consumption (gal.)			**Percentage Change from the Current Level**	
	(1)	(2)	(3)	(5)	(6)
	Current	Extra Fuel Tax \$0.25/gal.	Revenue Equivalent Product Tax	Extra Fuel Tax \$0.25/gal.	Revenue Equivalent Product Tax
Inelastic Demand for VMT					
Panel A: *Mean Miles vs. Heterogeneous Miles*					
IV RCLogit Mean Miles ($\gamma = 0.89$)	390.41	387.48	387.12	-0.75%	-0.84%
IV RCLogit EPD Miles ($\gamma = 1.01$)	380.18	374.97	376.77	-1.37%	-0.90%
Panel B: *Low Miles vs. High Miles in IV RCLogit EPD Miles Model*					
IV RCLogit EPD Miles (Low Miles)	252.15	251.10	250.77	-0.41%	-0.55%
IV RCLogit EPD Miles (High Miles)	508.21	498.82	502.77	-1.85%	-1.07%
The Elasticity of Demand for VMT Equals -0.77					
Panel C: *Mean Miles vs. Heterogeneous Miles*					
IV RCLogit Mean Miles ($\gamma = 0.89$)	390.41	381.75	387.26	-2.22%	-0.81%
IV RCLogit EPD Miles ($\gamma = 1.01$)	380.18	366.30	376.99	-3.65%	-0.84%
Panel D: *Low Miles vs. High Miles in IV RCLogit EPD Miles Model*					
IV RCLogit EPD Miles (Low Miles)	252.15	241.68	250.86	-4.15%	-0.51%
IV RCLogit EPD Miles (High Miles)	508.21	490.91	503.12	-3.40%	-1.00%

Notes: RCLogit stands for Random Coefficients Logit Model. EPD is short for Empirical Distribution. Consumers with expected annual vehicle miles traveled in the top half of the mileage distribution are in the "High Miles" group, and the rest are assigned to the "Low Miles" group. The product tax is calculated according to a car's per-mile fuel consumption. Averaging across all four scenarios, the difference between product taxes charged on cars with fuel economy ratings 30 MPG and 25 MPG is \$262.95 (in 2011 dollars).

2. IDENTIFYING THE ELASTICITY OF DRIVING: A FACTOR MODEL APPROACH AND IMPLICATIONS FOR USAGE-BASED TRANSPORTATION POLICIES

2.1 Introduction

Private vehicles dominate all categories of trips in the United States (Small et al., 2007). Therefore, personal vehicle transportation is of great importance to economic development and consumer welfare (Winston and Shirley, 1998). In the meantime, driving generates extensive energy use, substantial pollution and congestion (Parry et al., 2007). To address these externalities, state governments have been adjusting state motor fuel taxes, and the federal government has implemented a variety of measures to improve vehicle fuel efficiency.

A rise in fuel taxes intends to reduce driving by increasing the per-gallon fuel price. However, this effect might be canceled out by improved vehicle fuel efficiency that promoted under tightened fuel economy standards since an improvement in vehicle fuel efficiency reduces the per-mile fuel cost and hence may induce more driving. Recently, the idea of implementing a mileage tax has attracted policymakers' attention. Instead of charging for using motor fuel, this mileage tax explicitly increases the marginal cost of driving by charging drivers for their use of the road system by the amount that they drive (Langer et al., 2017). Overall, the real impact of these policies critically depends on consumer response in vehicle miles traveled (VMT) to changes in the variable driving cost engendered by policies (Gillingham et al., 2015).

In this paper, I focus on the intensive margin of consumer response to changes in fuel cost per mile, which is the primary component of variable driving cost counted by drivers (Hang et al., 2016). Specifically, I use a dataset constructed from vehicle registration and inspection records in Massachusetts over the period 2009-2014 to estimate the utilization elasticity of passenger vehicles, i.e., the elasticity of demand for VMT with respect to fuel cost per mile.

I apply the common correlated effects estimator developed by Pesaran (2006) to address unobserved common factors and regional heterogeneity in the data. Residual diagnostics confirm that the factor model setup does a better job of removing the cross-section dependence than standard fixed effects estimators do. The factor model result shows the elasticity of demand for driving with respect to fuel cost per mile is about -0.002. This estimate implies that a 10% increase in fuel cost per mile causes a 0.02% decrease in average daily miles traveled per vehicle, while results from fixed effects estimators suggest, when holding all else constant, the same percentage change in fuel cost per mile reduces driving by about 0.9%.

My fixed effects estimator results coincide with findings from the recent literature in which researchers construct large vehicle panels from odometer readings of extensive individual vehicles samples and apply various fixed effects controls for heterogeneity across drivers.[1] However, short-run elasticity estimates produced from such research designs are sensitive to the inclusion of controls for heterogeneity and whether or not vehicle characteristics are further subdivided when defining fixed effects (Hymel and Small, 2015).

A factor model setup developed by Pesaran (2006) is an ideal solution to capture unobserved heterogeneity in panels where the cross-section dimension is large. It allows for the presence of multiple unobserved common factors that correlate with the regressors. Further, the factor model setup lets factor loadings, which represent the effect of unobserved common factors, differ across observational units. Meanwhile, unlike the spatial econometric method, the factor model setup does not require prior knowledge about the structure of common factors and therefore lets the data de-

[1] Gillingham et al. (2015) builds a sample with 30,621,721 observations of 7,173,110 distinct vehicles from the Pennsylvania Vehicle Inspection Program. Knittel and Sandler (2013) applies a sample with 36,387,455 observations from the California Smog Check Program. Langer et al. (2017) uses a State Farm sample in Ohio with 228,910 driver-months observations.

termine the nature of unobserved heterogeneity (Totty, 2017).[2] Overall, relative to traditional fixed effects estimators, the factor model setup provides a more flexible way to address concerns about unobserved heterogeneity.

Given an inelastic demand for VMT with respect to fuel cost per mile, current policies leveraging the marginal cost of driving may not work as expected because of the low responsiveness. Instead, policymakers could consider options like congestion charges, parking charges, and other rights-based approaches to influence vehicle utilization and vehicle ownership.

The rest of the paper is laid out as follows. The next section reviews the relevant literature. Section 3 describes the data. In section 4, I introduce the empirical framework and state the identification strategy. Section 5 compares elasticity estimates from applying different estimators and presents residual diagnostic results. Section 6 concludes.

2.2 Related Literature

This paper contributes to existing literature studying consumers response on the intensive margin to changes in variable cost of driving, which remains a relevant as ever for policy analysis of price policies to reduce energy use, greenhouse gas emissions, and congestion from the extensive use of passenger vehicles in the US.

A review on older literature by Austin (2008) shows this estimated elasticity ranges from -0.10 to -0.16 in the short run. Small et al. (2007) estimate a system of equations capturing the simultaneous choice of aggregate VMT per capita, size of the automobile fleet, and fuel efficiency level of the fleet. Using data of US states from 1966 to 2001, the authors show that the short-run and long-run vehicle utilization elasticities are -0.045 and -0.222 respectively. Hymel et al. (2010) apply the same

[2] The spatial econometric method is the other approach to consider for dealing with unobserved cross-sectional dependence. Essentially, the spatial econometric method assumes that the structure of the cross-sectional correlation is related to location and distance among observational units. It requires a pre-specified connection or spatial weighting matrix to characterize the pattern of spatial dependence (Chudik and Pesaran, 2013).

simultaneous-equations method and estimates these two elasticities as -0.048 and -0.159 at sample average when using US data over 1984-2004. Overall, Small et al. (2007) summarize that in studies using aggregate data to investigate demand for travel in personal vehicles, utilization elasticities are usually between -0.1 and -0.3, with short-run elasticities typically smaller than long-run elasticities in absolute values.

Employing updated US data that covers the period 1996-2009, Hymel and Small (2015) show there is an upward shift in its magnitude for the short-run structural elasticity of VMT with respect to fuel cost per mile in the late 2000s when motor fuel prices were rapidly changing. This observation coincides with findings from recent studies making use of large scale odometer readings. As registration and inspection records of individual vehicles have become more accessible over the past few years, researchers construct panels with large cross-section dimension from this information and apply fixed effects estimators to investigate the vehicle utilization elasticity. Knittel and Sandler (2013) examine data from California's vehicle emissions testing program from 1996 to 2010 to obtain estimates of the VMT elasticity with respect to fuel cost per mile that vary between -0.04 and -0.26 depending on model specifications. Using vehicle inspection records from mandatory annual inspections in Pennsylvania from 2000 to 2010, Gillingham et al. (2015) show the one-year elasticity of VMT with respect to fuel cost per mile is about -0.097. More recently, Langer et al. (2017) apply a large monthly sample of drivers in Ohio from 2009 to 2011 to show that the range of the short-run elasticity of demand for automobile travel with respect to fuel cost per mile is from roughly -0.60 to slightly greater than zero depending on the accumulative vehicle odometer readings and whether sampled drivers live in rural area. All three studies employing odometer readings highlight the importance of accounting for heterogeneity across drivers and their vehicles when estimating vehicle utilization elasticities.

This paper is also associated with a growing empirical literature employing large linear panel data models to study economic activities across countries, regions, or industries as the availability of high-quality data has been increasing. Chudik and

Pesaran (2013) point out that in panels where both the cross-section dimension and the time series dimension are large, observational units need not be cross-sectionally independent even after conditioning on unit-specific regressors. The authors show that if the cross-sectional dependence correlates with regressors, conventional panel estimators such as fixed and random effects models can lead to misleading inference and even inconsistent estimates.

Pesaran (2006) proposes the Common Correlated Effects (CCE) estimation procedure, which augments standard pooled or mean group estimators with cross-sectional averages of all dependent and independent variables, to capture unobserved common factors that cause the cross-sectional dependence. This CCE estimation procedure has been gaining popularity in cross-sectional time series panel studies. For example, Eberhardt et al. (2013) employ this CCE estimation procedure to show that ignoring spillovers across different industries leads to bias in estimated private returns to R&D. Totty (2017) shows that minimum wage-employment elasticities produced from the factor model estimators are much smaller than the traditional Ordinary Least Square (OLS) results after the unobserved heterogeneity has been taken into consideration. Eberhardt and Teal (2017) apply the common factor model to capture the impact of heterogeneous technology in their cross-country analysis of the total factor productivity (TFP) while accounting for endogeneity and cross-section correlation that arise from global shocks.

2.3 Data

2.3.1 Data Sources

The primary data used in this paper is the municipality summary table from the Massachusetts Vehicle Census (Metropolitan Area Planning Council, 2016). The Massachusetts Vehicle Census (MAVC) is based on the Automated License and Registration System and a separate database containing records of vehicle inspections, both of which are administrative data sets maintained by the Massachusetts Registry of Motor Vehicles (Reardon et al., 2016).

In Massachusetts, passenger vehicle registration renewal is valid for two years while vehicles are required to be inspected annually and within seven days of sale. In the MAVC, odometer readings from vehicle inspection records and the dates of inspections are compared to calculate the mileage driven between inspections and the average daily mileage. This process leads to a series of mileage estimates for each vehicle with an estimated daily mileage during the intervening period. A registration period without a corresponding mileage estimate is retained but assigned a “false” value to indicate its invalidity.

The MAVC summary table compiles the information about individual vehicles into a variety of statistics for each municipality, such as total registered vehicles, average mileage per vehicle, average vehicle age of the automobile fleet, estimated fuel consumption and emissions. Since the MAVC constitutes a continuous longitudinal dataset, the municipality summary table is based on all the vehicle records that are valid on the median day of each calendar quarter. As the current version of MAVC includes approximately 34 million records for years from 2009 to 2014, there are 24 records for each municipality in the summary table.

For each vehicle included in the MAVC, the temporal overlap between the mileage estimate and the registration record is compared to the length of the mileage estimate period as a measure of data reliability. A high value for the percentage overlap means that the vehicle had the same owner and was garaged in the same location for a large portion of the mileage estimate period. The average daily mileage values employed in this paper are all valid estimates weighted by the percent overlap so that the least reliable estimates have relatively less impact on the resulting average.

The summary table also estimates an effective fuel efficiency level of the passenger vehicle fleet in each municipality based on aggregate mileage and fuel consumption of all vehicles with valid mileage estimates during the given sample period. This aggregate effective fuel efficiency level is calculated from the average per-mile fuel consumption of a vehicle weighted by the corresponding total miles traveled of that vehicle, which can be considered as the average fuel efficiency level of sampled vehicles weighted by utilization rates.

Since I am interested in the response in average daily mileage to the change in variable driving cost, I divide the average price of midgrade motor fuel in Massachusetts from the U.S. Energy Information Administration (EIA) by the MAVC effective fuel efficiency level to compute the aggregate per-mile fuel cost of driving for drivers in each town or city.[3] I include the average Massachusetts unemployment rate from the U.S. Bureau of Labor Statistics (BLS) and the total gross domestic product of Massachusetts from the U.S. Bureau of Economic Analysis (BEA) to control for changing macroeconomic conditions. Similarly, I bring in the national Consumer Confidence Index (CCI) from the Conference Board to account for consumersâĂŹ confidence towards the future economic condition. Finally, I employ the global price of Brent Crude from EIA as the instrumental variable for local fuel prices. I apply seasonal smooth to these four variables and convert all dollar-valued variables to 2010 dollars using the BLS Consumer Price Index (CPI). I also adjust all variables to the quarterly level to match the MAVC municipality summary table.

2.3.2 Data Description

In Massachusetts, representatives from local government, regional transit operators, and state transportation agencies form thirteen Metropolitan Planning Organizations (MPO) and Transportation Planning Organizations (TPO) to coordinate and manage transportation projects and programs that carried out in 351 towns and cities across the state.[4] I provide summary statistics of average daily miles traveled

[3] I also apply prices of the regular grade gasoline for robustness check. The results are nearly identical.
[4] Figure A.1 in the Appendix displays the geographic spread of all thirteen MPOs and TPOs in Massachusetts.

per vehicle, the effective vehicle fuel efficiency of the fleet, and average vehicle age of the fleet in table 2.1.[5]

In the first row of table 2.1, numbers in column (1) show Berkshire MPO covers 32 towns and cities which account for about 9.12% of total 351 municipalities. Statistics in column (2) of the same row tell that the average daily miles traveled per vehicle across all towns and cities included in Berkshire region has a mean value of 28.35 miles per day with the standard deviation equals 2.53. Similarly, the mean value of effective vehicle fuel efficiency level of the automobile fleet across all observational units and periods in Berkshire region is 19.64 miles per gallon (MPG), and the mean value of average vehicle age of the automobile fleet is 8.66 years.[6]

Table 2.1 shows that Boston region MPO covers almost one-third of all towns and cities in Massachusetts and vehicles in Greater Boston area are newer than vehicles in other parts of the state; vehicles in Montachusett region travel the most, while the vehicle fleet in Franklin region is the most fuel efficient. These numbers presented in table 2.1 are suggestive evidence of rich heterogeneity in the data. I expect the factor model setup to pick up such unobserved heterogeneity while not requiring any pre-specified structures.

I plot the evolution of fleet average fuel economy across all municipalities over time and the quarterly state average fuel price in Massachusetts from 2009 to 2014 in figure 2.1. Although the mean value of fleet average fuel efficiency has been increasing over the sample period, the improvement is modest. This makes sense because tightened fuel economy standards that went into effect in 2011 only affect new cars sold after that point while older vehicles take up the majority of the entire automobile fleet. Fuel economy standards have been criticized because they do not affect drivers decisions about VMT in their existing vehicles (Langer et al., 2017). Given a rapid change

[5]Although they are grouped by MPO and TPO in table 2.1, these key variables enter the regression analysis at the town and city average level, as they are reported in the MAVC municipality summary table.

[6]According to Bureau of Transportation Statistics, from 2008 to 2013, the average age of light vehicles in the U.S. increased by 12.2 percent because of a 40 percent drop in new vehicle sales due to the recession, while the increase in the average vehicle age was 3.5 percent over 2002-2007.

in the real fuel price in Massachusetts and a limited increase in fleet average fuel efficiency displayed in figure 2.1, the variation in the variable cost of driving that measured by dollars per mile mainly comes from changes in fuel prices.

In figure 2.2, I illustrate changes in the mean value of average daily miles traveled per vehicle across all municipalities alongside average fuel cost per mile that constructed from dividing the fuel price by the average fuel efficiency level of the fleet. While the response in VMT to changes in the variable cost of driving is obvious, I anticipate the factor model setup to account for both observed and unobserved common shocks flexibly.

2.4 Empirical Approach

2.4.1 Model Specification

The primary question asked in this paper is how changes in variable cost of driving affect decisions about VMT. Right now the major component of this variable driving cost is the per-mile fuel cost.[7] The group of consumers driving fuel-efficient vehicles will be less responsive to changes in fuel prices. To account for this fact, I divide the average per-gallon fuel price by the municipality-specific effective average vehicle fuel efficiency level of the automobile fleet to construct the variable driving cost, as measured by dollars per mile. This cost varies with both the fuel price and the average fuel efficiency level of the whole vehicle fleet in each town or city over time.

I model the demand for driving in municipality i during quarter t as a function of the variable driving cost at the fleet average level (P_{it}), the average vehicle age of the automobile fleet (A_{it}), Massachusetts unemployment rate (R_t), Massachusetts state

[7]Oregon has been experimenting the mileage tax by charging motorists and truckers for their use of the road system by the amount that they drive (Langer et al., 2017). California, Hawaii, the state of Washington, Connecticut, Delaware, New Hampshire, and Pennsylvania are expected to conduct such test.

GDP (G_t), national CCI (C_t), quarter-of-the-year fixed effects (λ_t), and municipality fixed effects (ψ_i): $VMT_{it} = f(P_{it}, A_{it}, R_t, G_t, I_t, \lambda_t, \psi_i)$. I further specify it as

$$v_{it} = \beta_i p_{it} + \gamma_i a_{it} + \alpha_r r_t + \alpha_g g_t + \alpha_c c_t + \lambda_t + \psi_i + \epsilon_{it} \tag{2.1}$$

in which ϵ_{it} is a mean-zero stochastic error term and lower case letters denote logarithms of VMT and inputs that drives the demand for VMT.[8] For generality, I assume $\beta_i = \beta + h_i$ and $\gamma_i = \gamma + k_i$ where both h_i and k_i are random variables that induces variation of parameters across observational units (Greene, 2011). Since the MAVC summary table consists of data collected from 351 municipalities in Massachusetts over five years, I interpret the β_i estimate in this log-linear estimating equation as a mid-run elasticity of demand for driving with respect to the variable driving cost.

2.4.2 Identification

The primary interest of this paper is to estimate the price elasticity of VMT at the aggregated municipality level. Drivers' differential responses to changes in fuel cost per mile help identify the parameter of interest. A biased estimate of β_i would arise from omitted variables that are correlated with fuel prices and affect VMT through the effective fuel efficiency level of the vehicle fleet in each municipality.[9]

In equation 2.1, the municipality fixed effects capture unobserved local characteristics that correlated with observed influences on driving demand. The average vehicle age captures the impact of unobserved fleet characteristics on demand for VMT. Also, contemporaneous macroeconomic and seasonal conditions could affect how much people drive. Thus I control for that potential source of bias by including state level macroeconomic variables and quarter-of-the-year fixed effects. Further-

[8]This constant elasticity model is one of the most frequently estimated in the literature. It can be derived from a model in which an individual has Cobb-Douglas preferences over VMT and all other goods (Linn, 2016).

[9]Small et al. (2007) suggests that both the evolution of vehicle fuel efficiency itself and congestion could be endogenous factors that influence the demand for VMT.

more, I include national CCI to control for the change in consumer perspective as the economy was recovering from the dramatic economic downturn beginning in 2008.

In addition to including a rich set of fixed effects to deal with a variety of possible confounding factors, I apply instrumental variables to address the impact of localized VMT demand shocks on fuel prices. The fundamental concern is that unobserved local shocks may shift the demand for driving outside, which will increase the fuel price in equilibrium and lead to a supply response such as immediately refining more motor fuel and moving it to Massachusetts (Gillingham et al., 2015). Since the crude oil price is unquestionably the primary determinant of the local motor fuel price, it can be expected to be a strong instrument. Therefore, I instrument for the Massachusetts fuel price with the global price of Brent Crude from EIA. It seems reasonable that the only way Brent Crude prices influence the demand for driving in Massachusetts is through local motor fuel prices, so the exclusion restriction holds (Gillingham, 2014).

When employing cross-sectional time series to study aggregate travel demand, researchers may choose to apply pooled regression, fixed effects or random effects estimators depending on the identification assumption (Small et al., 2007).[10] When such models are applied, we expect time-specific variables and time dummies to purge correlations across observational units that arise from common shocks (Eberhardt et al., 2013). Therefore, if the model is correctly specified, we would expect well-behaved, serially uncorrelated, stationary, and most importantly, cross-sectionally independent regression residuals $\hat{\epsilon}_{it}$ (Greene, 2011).

However, it has been recognized that in panels with large cross-section dimension, even after conditioning on unit-specific regressors, individual observational units may not be cross-sectionally independent. Cross-section correlations of errors could come from omitted common effects or unobserved interactions within socioeconomic networks. In complex models, interdependence could also arise from common correlated

[10] Voith (1997) and Petitte (2001) are examples of early studies using cross-sectional time series to study the urban transportation. This type of research design is also adopted in the literature examining aggregate demand for motor gasoline, in which long panels of U.S. states are constructed for analysis.

reaction of observational units to some external events. Moreover, if the source generating the cross-sectional dependence correlates with regressors, conventional panel estimators may fail to yield correct inference and consistent estimates (Chudik and Pesaran, 2013).

In a simplified version of equation 2.1

$$v_{it} = \beta_i x_{it} + \epsilon_{it}, \tag{2.2}$$

I let the cross-sectional dependence arise from the shared common factor f_t in error structure ϵ_{it} and in input x_{it} such that

$$\epsilon_{it} = \delta_i f_t + \psi_i + \nu_{it} \tag{2.2a}$$

$$x_{it} = \sigma_i f_t + \lambda_t + \phi_i + \mu_{it} \tag{2.2b}$$

in which ν_{it} and μ_{it} are stochastic shocks. In this setup, factor loading δ_i and σ_i trace the strength of unobserved common factor f_t that drives VMT as well as inputs generating the demand for driving. However, if neither δ_i and σ_i is zero on average, the estimate of β_i from equation 2.2 will be biased and inconsistent as described in this following equation

$$v_{it} = (\beta_i + \delta_i \sigma_i^{-1}) x_{it} + (\psi_i - \delta_i \sigma_i^{-1} \phi_i) + (\nu_{it} - \delta_i \sigma_i^{-1} \mu_{it} - \delta_i \sigma_i^{-1} \lambda_t) \tag{2.3}$$

in which $E[\beta_i + \delta_i \sigma_i^{-1}] \neq \beta_i$ if $\delta_i \sigma_i^{-1} \neq 0$.

2.4.3 Empirical Implementation

To account for the impact of unobserved common factors that may differ across towns and cities as illustrated in equations 2.2a and 2.2b, I apply the Pesaran (2006) Common Correlated Effects Mean Group (CCEMG) estimator. This common factor approach assumes that the error term, as well as regressors contain a finite number of unobserved common factors and their influences are allowed to differ across different observational units. This setup is particularly useful when analyzing travel demand using large panels like the MAVC municipality summary table because it is hard to exhaust elements that affect the demand for driving or specify the structure of correlations among different locations. Further, the cross-sectional dependence may result from events with heterogeneous impacts across municipalities, such as temporary interstate closures that affect a limited group of towns and cities, or some extreme weather conditions and natural disasters that have a significant impact across the whole state but affect different regions with various intensities.

Pesaran (2006) shows that, when the cross-section dimension is large, under some general conditions, unobserved common factors can be captured by adding cross-sectional averages of both dependent and independent variables as additional regressors. To show the intuition of CCEMG, I combine equation 2.2 and equation 2.2a to get

$$v_{it} = \beta_i x_{it} + \delta_i f_t + \psi_i + \nu_{it}. \tag{2.4}$$

Next I take the cross-sectional average of equation 2.4 to obtain an expression of the unobserved common factor by rearranging terms

$$\begin{aligned} & \bar{v}_t = \bar{\beta}\bar{x}_t + \bar{\delta} f_t + \bar{\psi} \quad \text{given } \bar{\nu}_{it} \to 0 \text{ as } N \to \infty \\ \iff \quad & f_t = \bar{\delta}^{-1}(\bar{v}_t - \bar{\beta}\bar{x}_t - \bar{\psi}) \end{aligned} \tag{2.5}$$

in which cross-sectional averages at time t are defined as $\bar{v}_t = N^{-1}\sum_{i=1}^{N} v_{it}$ and $\bar{x}_t = N^{-1}\sum_{i=1}^{N} x_{it}$. Substituting for f_t in equation 2.4 gives

$$\begin{aligned} & v_{it} = \beta_i x_{it} + \delta_i \bar{\delta}^{-1}(\bar{v}_t - \bar{\beta}\bar{x}_t - \bar{\psi}) + \psi_i + \nu_{it} \\ \iff \quad & v_{it} = \beta_i x_{it} + \pi_{1i}\bar{v}_t + \pi_{2i}\bar{x}_t + \pi_{3i} + \nu_{it}. \end{aligned} \tag{2.6}$$

The advantage of this CCEMG approach is that it does not require prior knowledge of the number or the structure of unobserved common factors (Chudik and Pesaran, 2013). The focus of this estimation approach is to obtain an unbiased estimate for the mean of heterogeneous β_i. In equation 2.6, π_{1i}, π_{2i} and π_{3i} are municipality-specific parameters so that they are able to capture the heterogeneity in factor loading δ_i. However, these parameters are not interpretable because they contain various averages of unknowns and therefore should be only seen as components accounting for the cross-section dependence in the data (Eberhardt et al., 2013).

2.5 Results

In this section, I report empirical results from analyzing the VMT data from 351 municipalities in Massachusetts over 24 quarters from 2009 to 2014. I apply the specification described in equation 2.1 while employing a number of estimators differ in their assumptions about cross-section dependence and unobserved common factors. I use residual cross-section correlation tests and residual stationarity tests developed in Pesaran (2015) and Pesaran (2007) to evaluate rival estimators. Further, I use the root mean squared error (RMSE) statistic to measure the goodness of fit for each regression model.

In column (1) of table 2.2, I pool together 8,424 observations and apply the OLS method to estimate equation 2.1. In terms of partial elasticity, when holding all else constant, the estimate of β indicates that a 10% increase in the variable driving cost (in dollars per mile) leads to more than 0.8% decrease in average daily miles traveled per vehicle. In the meantime, drivers put less mileage on older vehicles. Impacts of other macroeconomic conditions considered in equation 2.1 have sensible signs but are not significant at 10% level.

Since I concern about the possible scenario in which localized VMT demand shocks shift motor fuel prices, I instrument for the Massachusetts fuel price with the global price of Brent Crude from EIA in estimation model (2) through (6).[11] In column (2), results from the pooled OLS estimator with an instrumental variable (IV) have the similar pattern as those from pooled OLS estimator without an IV, while the price elasticity of demand for driving is slightly smaller in absolute value.

Columns (3) and (4) present results from employing fixed effects estimators with IVs. The parameter estimate from a two-way fixed effects estimator, in which both municipality fixed effects and quarter-of-the-year fixed effects are applied, suggests a more elastic demand for driving relative to results from OLS estimators. These two fixed effects estimators produce comparable estimates for the impact of average vehicle age on vehicle utilization. Meanwhile, both estimators yield negative and significant estimates for the impact of unemployment on VMT, which indicates that vehicle owners drive less when the unemployment rate is high. Also, parameter estimates from the two-way fixed effect estimator suggest that consumers drive more during the second and the third quarter, which makes sense as people tend to drive more starting from the end of spring through late summer because of increased leisure activities. However, estimates from fixed effects models disagree with those from pooled OLS estimators on the impact of state GDP and national consumer perspective.

While both pooled OLS and fixed effects estimators assume common parameters for regressors in equation 2.1 across all municipalities, mean group type estimators examined in column (5) and (6) allow for heterogeneity by running municipality-specific regressions and then averaging coefficients across the panel (Eberhardt, 2012).[12] In column (5) of table 2.2, I report results from applying the mean group (MG) estimator proposed by Pesaran and Smith (1995). The pattern of estimates from the MG estimator is very similar to that from the two-way fixed effects estimator, except that the magnitude is larger for the effect of average vehicle age on demand for driving.

[11] This Brent Crude price is a strong instrument as the t-static is over 700 for the logarithm of the crude price and the F-statistic is 2.3×10^5 in the first stage regression. The first stage results are available in the Appendix.

[12] Boyd and Smith (2002) suggests that mean group results for individual observational unit are difficult to interpret and not reliable unless the time series dimension is large, while panel averages establish a reliable mean estimate.

Moving on to column (6), the CCEMG estimator produces a negative and significant estimate for the variable driving cost but with a much smaller magnitude when comparing to results from previous regressions. It indicates that, when holding all else constant, a 10% increase in the per-mile fuel cost decreases the average daily miles traveled per vehicle by only about 0.02%. Meanwhile, results from both mean group type estimators suggest people don't drive older vehicles as much as they do to newer ones, which may come down to the capital depreciation or a higher on-going maintenance cost for older vehicles. However, none of the parameter estimates for three macroeconomic condition variables are significant at the 10% level when this CCEMG estimator is applied.

Turning to the post-estimation diagnostic presented at the bottom of table 2.2, CCEMG estimator is the only case that not suffering from cross-sectional correlated residuals while both mean group type estimators generate stationary residual sequences.[13] It seems that the standard pooled OLS and fixed effects estimators are seriously misspecified when both cross-section and time series dimensions are large in the data. The diagnostic result indicates that the CCEMG estimator is the preferred setup to apply when analyzing the Massachusetts municipality summary table. The data shows that, in Massachusetts, the response in average daily miles traveled per vehicle across different towns and cities to changes in fuel cost per mile is significant but negligible.

Results reported in table 2.2 are based on equation 2.1, which models VMT as a function of fuel cost per mile. It is also natural to consider how consumers change the amount they drive in response to changing retail fuel prices. To examine the elasticity of VMT with respect to the per-gallon fuel price, I estimate equation 2.1 using the quarterly state average price of midgrade motor fuel in Massachusetts. By replacing the per-mile fuel cost with this per-gallon fuel cost, I remove the effective average fuel efficiency level of the vehicle fleet that varies across municipalities. With

[13] The Pesaran (2015) cross-section dependence test is based on the average of pair-wise correlations. Using a standard normal distribution, the null hypothesis of the test is that there is only weak cross-section dependence, while the alternative is that the cross-section dependence is strong.

this modification, I entirely rely on the variation in state average fuel prices over time to identify the elasticity of driving.

I estimate equation 2.1 with the per-gallon fuel cost using different estimators and report results in table 2.3. The pattern of estimated elasticities in regression models with IVs is almost identical to that shown in table 2.2.[14] The CCEMG result suggests that a 10% increase in fuel cost per gallon causes about a 0.03% decrease in the average daily miles traveled per vehicle. This result from regressions using retail fuel prices provides further evidence suggestive of consumers' low responsiveness to changes in variable driving cost.

2.6 Concluding Remarks

In this paper, I have shown that accounting for unobserved common factors and regional heterogeneity is critical when estimating the elasticity of VMT with respect to the variable driving cost in panels where both cross-section and time series dimensions are large. Cross-sectional dependence tests support the presence of common factors in residuals produced from applying either standard fixed effects estimators or traditional mean group estimators. Therefore, significant and substantial effects of fuel cost per mile on passenger vehicle utilization obtained from applying fixed effects estimators to large vehicle panels in recent literature require further inspection if unobserved common factors in the data may cause the bias.

Given the MAVC municipality summary table employed in this paper, CCEMG estimator developed by Pesaran (2006) is the preferred setup to use according to post-estimation tests. CCEMG results indicate that a 10% increase in fuel cost per mile causes only a 0.02% decrease in average daily miles traveled per vehicle, while results from applying fixed effects estimators suggest that the response in vehicle utilization to the same change in variable cost of driving can be as large as 0.9%

[14]Both Frondel and Vance (2013) and Gillingham et al. (2015) find little difference in estimated vehicle utilization elasticities regardless of choosing fuel cost per mile or fuel price per gallon as the key independent variable.

when holding all else constant. If consumers indeed barely respond to changes in the marginal cost of driving that engendered by current usage-based policies, rights-based approaches (e.g., congestion charges) might be reasonable alternatives to influence vehicle utilization and vehicle ownership.

Figures And Tables

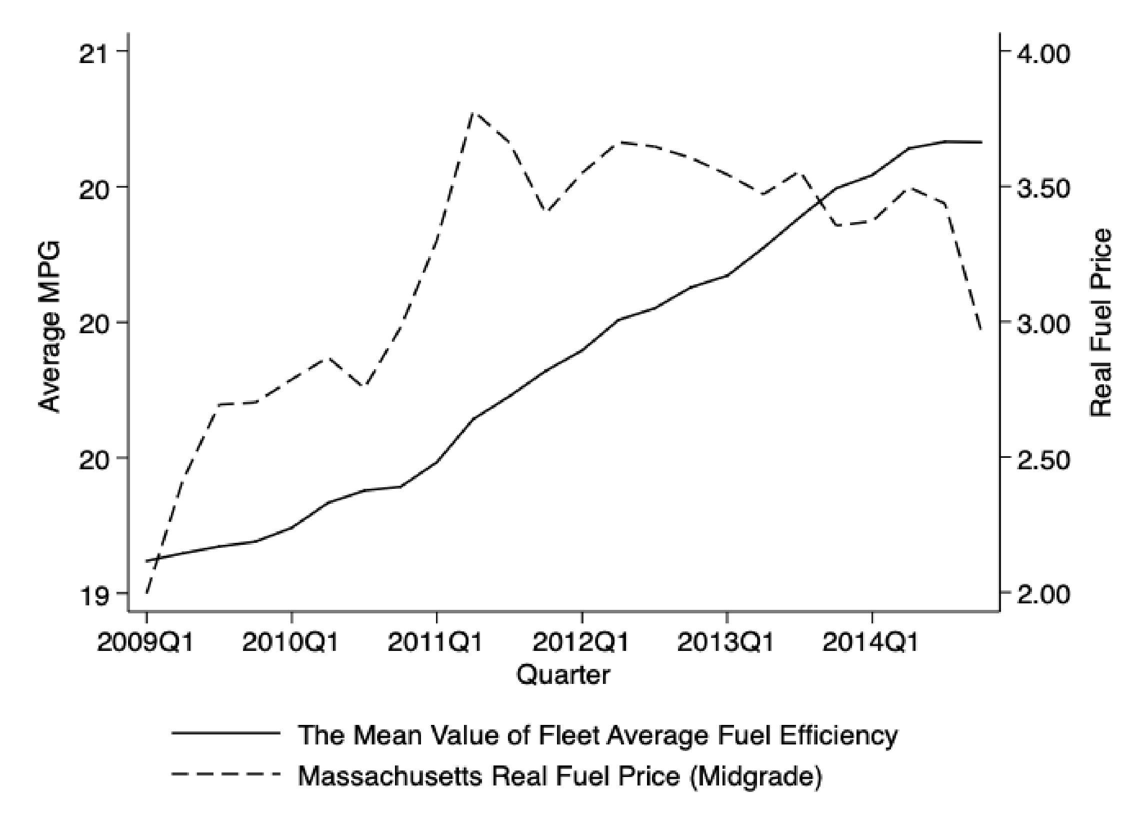

Figure 2.1. MPG stands for miles per gallon. Prices are normalized to 2010 dollars. Fuel Prices Data Source: EIA Monthly Motor Fuel Prices Table.

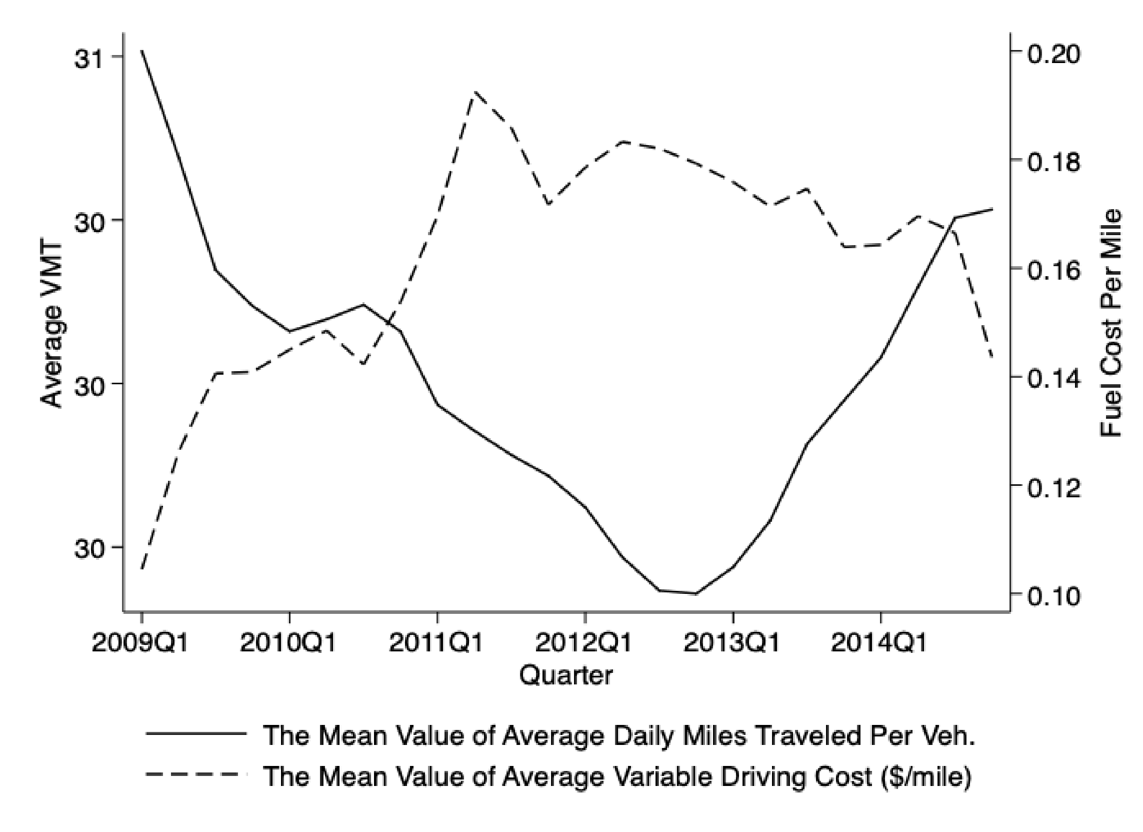

Figure 2.2. VMT stands for vehicle miles traveled. Prices are normalized to 2010 dollars. Fuel Prices Data Source: EIA Monthly Motor Fuel Prices Table.

Table 2.1.
Summary Statistics by Regions

	(1) Muni. Count		(2) Daily Mileage		(3) Effective MPG		(4) Fleet Veh. Age	
MPO/TPO	Freq.	Percent	Mean	S.D.	Mean	S.D.	Mean	S.D.
Berkshire MPO	32	9.12	28.35	2.53	19.64	0.75	8.66	0.55
Boston MPO	97	27.64	28.14	3.19	20.05	0.84	8.14	0.63
Cape Cod MPO	15	4.27	27.97	3.08	19.45	0.58	8.95	0.49
Central MA MPO	40	11.40	32.31	1.88	19.88	0.67	8.34	0.64
Franklin TPO	26	7.41	32.31	2.16	20.32	0.79	9.35	0.63
Martha's Vineyard MPO	7	1.99	20.76	2.46	18.67	0.73	10.45	0.78
Merrimack Valley MPO	15	4.27	32.81	2.19	19.88	0.62	8.36	0.86
Montachusett MPO	22	6.27	33.61	2.08	20.01	0.67	8.59	0.57
Nantucket MPO	1	0.28	14.06	0.24	17.05	0.29	9.93	0.66
Northern Middlesex MPO	9	2.56	30.93	2.16	19.89	0.65	8.29	0.62
Old Colony MPO	17	4.84	31.13	2.35	19.32	0.64	8.56	0.62
Pioneer Valley MPO	43	12.25	30.19	3.03	19.85	0.91	9.01	0.59
Southeastern MA MPO	27	7.69	31.47	2.92	19.77	0.61	8.85	0.67
Total	351	100	30.01	3.68	19.86	0.82	8.61	0.78

Notes: MPO represents the Metropolitan Planning Organization, TPO stands for Transportation Planning Organization. Municipality Count Frequency column shows how many towns and cities are included in a corresponding MPO or TPO. Mean Daily Mileage column tells the mean value of average daily miles traveled per vehicle. Mean Effective MPG column reports the mean value of effective vehicle fuel efficiency of the fleet, as measured by miles per gallon (MPG). Mean Fleet Vehicle Age column describes the mean value of average vehicle age of the fleet.

Table 2.2.
Regression Results for Variable Driving Cost

	POLS (1)	**POLS,IV (2)**	**Muni. FE,IV (3)**	**2FE,IV (4)**	**MG,IV (5)**	**CCEMG,IV (6)**
ln(dollar per mile)	-0.0862***	-0.0732***	-0.0792***	-0.0894***	-0.0748***	-0.0024**
	(0.0170)	(0.0169)	(0.0021)	(0.0027)	(0.0021)	(0.0015)
ln(vehicle age)	-0.3650***	-0.3660***	-0.0696**	-0.0699**	-0.1830***	-0.1110***
	(0.0223)	(0.0230)	(0.0278)	(0.0298)	(0.0181)	(0.0309)
ln(MA unemployment)	-0.0420	-0.0608	-0.0629***	-0.0549***	-0.0625***	-0.0005
	(0.0396)	(0.0395)	(0.0037)	(0.0056)	(0.0037)	(0.0039)
ln(MA GDP)	0.0190	0.1150	-0.0445**	0.0047	0.0164	-0.0037
	(0.1300)	(0.1270)	(0.0178)	(0.0259)	(0.0130)	(0.0140)
ln(national CCI)	0.1040	0.2750	0.0298	-0.1380***	-0.1190***	0.0027
	(0.3670)	(0.3690)	(0.0303)	(0.0433)	(0.0267)	(0.0175)
2nd Quarter				0.0095***		
				(0.0003)		
3rd Quarter				0.0039***		
				(0.0003)		
4th Quarter				0.0007		
				(0.0004)		
Diagnostic:						
$\hat{\epsilon}$ CD Test *p-value*	< 0.01	<0.01	<0.01	<0.01	<0.01	0.14
$\hat{\epsilon}$ I(1) *p-value*	1.00	1.00	1.00	1.00	<0.01	<0.01
RMSE	0.1269	0.1270	0.0148	0.0147	0.0106	0.0065

Notes: The dependent variable is ln(average daily miles traveled per vehicle). Regressions are for N = 351 municipalities, n = 8,424 observations in the levels specifications. For estimators (1) through (4), values in parentheses are robust standard errors accommodate for heteroskedasticity and autocorrelation; standard errors in parentheses for estimator (5) and (6) follow Pesaran and Smith (1995) in which a z-test is applied. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Intercept estimates and average estimates on cross-section average of CCEMG are omitted. Pesaran (2015) Residual Cross-sectional Dependence Test (i.e., CD Test) has the null hypothesis that there is only weak cross-section dependence, while the alternative is that the cross-section dependence is strong. Residual Order of Integration Test reports results for a Pesaran (2007) Panel Unit Root Test with 2 lags and a null of nonstationarity (i.e., I(1)). RMSE is the root mean squared error.

Table 2.3.
Regression Results for Retail Gas Price

	POLS (1)	**POLS,IV (2)**	**Muni. FE,IV (3)**	**2FE,IV (4)**	**MG,IV (5)**	**CCEMG,IV (6)**
ln(dollar per gallon)	-0.0678***	-0.0736***	-0.0799***	-0.0914***	-0.0752***	-0.0027**
	(0.0162)	(0.0170)	(0.0021)	(0.0027)	(0.0021)	(0.0015)
ln(vehicle age)	-0.3710***	-0.3710***	-0.0623**	-0.0636**	-0.1930***	-0.1120***
	(0.0230)	(0.0230)	(0.0278)	(0.0298)	(0.0180)	(0.0309)
ln(MA unemployment)	-0.0712*	-0.0630	-0.0654***	-0.0550***	-0.0646***	-0.0006
	(0.0385)	(0.0392)	(0.0037)	(0.0056)	(0.0038)	(0.0039)
ln(MA GDP)	0.1070	0.1420	-0.0218	0.0419	0.0465***	-0.0032
	(0.1280)	(0.1320)	(0.0178)	(0.0263)	(0.0130)	(0.0140)
ln(national CCI)	0.3650	0.2900	0.0355	-0.1550***	0.1370***	0.0200
	(0.3610)	(0.3670)	(0.0303)	(0.0435)	(0.0268)	(0.0212)
2nd Quarter				0.0097***		
				(0.0003)		
3rd Quarter				0.0037***		
				(0.0003)		
4th Quarter				0.0003		
				(0.0004)		
Diagnostic:						
$\hat{\epsilon}$ CD Test *p-value*	<0.01	<0.01	<0.01	<0.01	<0.01	0.14
$\hat{\epsilon}$ I(1) *p-value*	1.00	1.00	1.00	1.00	0.00	0.00
RMSE	0.1270	0.1270	0.0148	0.0147	0.0106	0.0065

Notes: The dependent variable is ln(average daily miles traveled per vehicle). Regressions are for N = 351 municipalities, n = 8,424 observations in the levels specifications. For estimators (1) through (4), values in parentheses are robust standard errors accommodate for heteroskedasticity and autocorrelation; standard errors in parentheses for estimator (5) and (6) follow Pesaran and Smith (1995) in which a z-test is applied. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Intercept estimates and average estimates on cross-section average of CCEMG are omitted. Pesaran (2015) Residual Cross-sectional Dependence Test (i.e., CD Test) has the null hypothesis that there is only weak cross-section dependence, while the alternative is that the cross-section dependence is strong. Residual Order of Integration Test reports results for a Pesaran (2007) Panel Unit Root Test with 2 lags and a null of nonstationarity (i.e., I(1)). RMSE is the root mean squared error.

REFERENCES

REFERENCES

Ackerberg, D., C. L. Benkard, S. Berry, and A. Pakes (2007). Econometric tools for analyzing market outcomes. *Handbook of econometrics 6*, 4171–4276.

Allcott, H. and M. Greenstone (2012). Is there an energy efficiency gap? *The Journal of Economic Perspectives 26*(1), 3–28.

Allcott, H., S. Mullainathan, and D. Taubinsky (2014). Energy policy with externalities and internalities. *Journal of Public Economics 112*, 72–88.

Allcott, H. and N. Wozny (2014). Gasoline prices, fuel economy, and the energy paradox. *Review of Economics and Statistics 96*(5), 779–795.

Anderson, S. T., R. Kellogg, J. M. Sallee, and R. T. Curtin (2011). Forecasting gasoline prices using consumer surveys. *American Economic Review 101*(3), 110–14.

Austin, D. (2008). *Effects of gasoline prices on driving behavior and vehicle markets*. Government Printing Office.

Baumeister, C. and L. Kilian (2016). Forty years of oil price fluctuations: Why the price of oil may still surprise us. *Journal of Economic Perspectives 30*(1), 139–60.

Bento, A. M., L. H. Goulder, M. R. Jacobsen, and R. H. Von Haefen (2009). Distributional and efficiency impacts of increased us gasoline taxes. *American Economic Review 99*(3), 667–99.

Bento, A. M., S. Li, and K. Roth (2012). Is there an energy paradox in fuel economy? a note on the role of consumer heterogeneity and sorting bias. *Economics Letters 115*(1), 44–48.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of political Economy 112*(1), 68–105.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Boyd, D. and R. P. Smith (2002). Some econometric issues in measuring the monetary transmission mechanism, with an application to developing countries. *Monetary transmission in diverse economies*, 68–99.

Busse, M. R., C. R. Knittel, J. Silva-Risso, and F. Zettelmeyer (2016). Who is exposed to gas prices? how gasoline prices affect automobile manufacturers and dealerships. *Quantitative Marketing and Economics 14*(1), 41–95.

Busse, M. R., C. R. Knittel, and F. Zettelmeyer (2013). Are consumers myopic? evidence from new and used car purchases. *The American Economic Review 103*(1), 220–256.

Chudik, A. and M. H. Pesaran (2013). Large panel data models with cross-sectional dependence: a survey. *CAFE Research Paper* (13.15).

Copeland, A., W. Dunn, and G. Hall (2011). Inventories and the automobile market. *The RAND Journal of Economics 42*(1), 121–149.

Davis, L. W. and L. Kilian (2011). Estimating the effect of a gasoline tax on carbon emissions. *Journal of Applied Econometrics 26*(7), 1187–1214.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic literature 47*(2), 315–72.

Eberhardt, M. (2012). Estimating panel time-series models with heterogeneous slopes. *The Stata Journal 12*(1), 61–71.

Eberhardt, M., C. Helmers, and H. Strauss (2013). Do spillovers matter when estimating private returns to r&d? *Review of Economics and Statistics 95*(2), 436–448.

Eberhardt, M. and F. Teal (2017). The magnitude of the task ahead: Macro implications of heterogeneous technology. *Review of Income and Wealth*.

EPA (2018). Inventory of u.s. greenhouse gas emissions and sinks: 1990-2016. *Working Paper*, ES–12.

EPA and NHTSA (2010). Light-duty vehicle greenhouse gas emission standards and corporate average fuel economy standards; final rule.

Frondel, M. and C. Vance (2013). Re-identifying the rebound: what about asymmetry? *The Energy Journal*, 43–54.

Fullerton, D. and S. E. West (2002). Can taxes on cars and on gasoline mimic an unavailable tax on emissions? *Journal of Environmental Economics and Management 43*(1), 135–157.

Fullerton, D. and S. E. West (2010). Tax and subsidy combinations for the control of car pollution. *The BE Journal of Economic Analysis & Policy 10*(1).

Gandhi, A. and J.-F. Houde (2016). Measuring substitution patterns in differentiated products industries. *University of Wisconsin-Madison and Wharton School*.

Gillingham, K. (2014). Identifying the elasticity of driving: evidence from a gasoline price shock in california. *Regional Science and Urban Economics 47*, 13–24.

Gillingham, K., A. Jenn, and I. M. Azevedo (2015). Heterogeneity in the response to gasoline prices: Evidence from pennsylvania and implications for the rebound effect. *Energy Economics 52*, S41–S52.

Gillingham, K., D. Rapson, and G. Wagner (2016). The rebound effect and energy efficiency policy. *Review of Environmental Economics and Policy 10*(1), 68–88.

Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the us automobile industry. *Econometrica: Journal of the Econometric Society*, 891–951.

Greene, W. (2011). *Econometric Analysis*. Pearson Education.

Grigolon, L., M. Reynaert, and F. Verboven (2018, August). Consumer valuation of fuel costs and tax policy: Evidence from the european car market. *American Economic Journal: Economic Policy 10*(3), 193–225.

Hang, D., D. McFadden, K. Train, and K. Wise (2016). Is vehicle depreciation a component of marginal travel cost?: A literature review and empirical analysis. *Journal of Transport Economics and Policy (JTEP) 50*(2), 132–150.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hassett, K. A. and G. E. Metcalf (1993). Energy conservation investment: Do consumers discount the future correctly? *Energy Policy 21*(6), 710–716.

Hausman, J. A. (1979). Individual discount rates and the purchase and utilization of energy-using durables. *The Bell Journal of Economics*, 33–54.

Hymel, K. M. and K. A. Small (2015). The rebound effect for automobile travel: asymmetric response to price changes and novel features of the 2000s. *Energy Economics 49*, 93–103.

Hymel, K. M., K. A. Small, and K. Van Dender (2010). Induced demand and rebound effects in road transport. *Transportation Research Part B: Methodological 44*(10), 1220–1241.

Innes, R. (1996). Regulating automobile pollution under certainty, competition, and imperfect information. *Journal of Environmental Economics and Management 31*(2), 219–239.

Kilian, L. and E. R. Sims (2006). The effects of real gasoline prices on automobile demand: a structural analysis using micro data. *University of Michigan, manuscript*.

Knittel, C. R., D. L. Miller, and N. J. Sanders (2016). Caution, drivers! children present: Traffic, pollution, and infant health. *Review of Economics and Statistics 98*(2), 350–366.

Knittel, C. R. and R. Sandler (2013). The welfare impact of indirect pigouvian taxation: Evidence from transportation. Technical report, National Bureau of Economic Research.

Langer, A., V. Maheshri, and C. Winston (2017). From gallons to miles: A disaggregate analysis of automobile travel and externality taxes. *Journal of Public Economics 152*, 34–46.

Leard, B., J. Linn, and Y. C. Zhou (2017). How much do consumers value fuel economy and performance? *Evidence from technology adoption.(RFF Policy Report)*.

Linn, J. (2016). The rebound effect for passenger vehicles. *Energy Journal 37*(2).

Lu, S. (2006). Vehicle survivability and travel mileage schedules. Technical report, National Highway Traffic Safety Administration.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of economics & management strategy 9*(4), 513–548.

Pakes, A. (1994). Dynamic structural models, problems and prospects. Advances in Econometrics. Sixth World Congress, Cambridge University Press.

Parry, I. W., M. Walls, and W. Harrington (2007). Automobile externalities and policies. *Journal of economic literature 45*(2), 373–399.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica 74*(4), 967–1012.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of applied econometrics 22*(2), 265–312.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric Reviews 34*(6-10), 1089–1117.

Pesaran, M. H. and R. Smith (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics 68*(1), 79–113.

Petitte, R. A. (2001). Fare variable construction and rail transit ridership elasticities: case study of the washington, dc, metrorail system. *Transportation research record 1753*(1), 102–110.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of political Economy 110*(4), 705–729.

Reardon, T., E. Irvin, S. Brunton, M. Hari, P. Reim, and K. Gillingham (2016). Quantifying vehicle miles traveled from motor vehicle inspection data: The Massachusetts vehicle census. In *Transportation Research Board 95th Annual Meeting*, Number 16-0634.

Simeonova, E., J. Currie, P. Nilsson, and R. Walker (2018). Congestion pricing, air pollution and childrenâĂŹs health. Technical report, National Bureau of Economic Research.

Small, K. A., K. Van Dender, et al. (2007). Fuel efficiency and motor vehicle travel: the declining rebound effect. *ENERGY JOURNAL-CAMBRIDGE MA THEN CLEVELAND OH- 28*(1), 25.

Small, K. A., E. T. Verhoef, and R. Lindsey (2007). *The economics of urban transportation.* Routledge.

Stock, J. H. and M. Yogo (2002). Testing for weak instruments in linear iv regression.

Totty, E. (2017). The effect of minimum wages on employment: A factor model approach. *Economic Inquiry 55*(4), 1712–1737.

Train, K. E. (2009). *Discrete choice methods with simulation.* Cambridge university press.

Train, K. E. and C. Winston (2007). Vehicle choice behavior and the declining market share of us automakers. *International economic review 48*(4), 1469–1496.

Voith, R. (1997). Fares, service levels, and demographics: what determines commuter rail ridership in the long run? *Journal of Urban Economics 41*(2), 176–197.

Whitefoot, K., M. Fowlie, and S. Skerlos (2011). Product design response to industrial policy: Evaluating fuel economy standards using an engineering model of endogenous product design. *Energy Institute at Haas Working Paper No. WP-214*.

Winston, C. and C. Shirley (1998). *Alternate route: Toward efficient urban transportation.* Brookings Institution Press.

APPENDICES

A. APPENDIX: RESPONSE OF NEW CAR BUYERS TO ALTERNATIVE ENERGY POLICIES

A.1 Consumer Response to Policy Changes

Plugging new prices into equation 1.5, the deterministic part of the conditional indirect utility of consumer i for car model v in market t is now

$$\phi_{ivt} \equiv \alpha_i \; x_v + \beta_i \; (p_v + \frac{\tau^{mpg}}{mpg_v} + \gamma \; \rho \; \frac{m_{it}}{mpg_v} \; (E[g_i] + \tau^g)) + \zeta_{vt}.$$

The probability of consumer i choosing model v in market t becomes

$$s_{ivt} = \frac{exp(\phi_{ivt})}{1 + \sum_{v'=1}^{J} exp(\phi_{iv't})}.$$

To inspect the consumer response to the implementation of a new gasoline tax and a product tax, I differentiate the individual choice probability with respect to both τ^g and τ^{mpg}. Consumer i's response given the change in gasoline tax is

$$\begin{aligned} \frac{\partial s_{ivt}}{\partial \tau^g} &= s_{ivt} \frac{\partial \phi_{ivt}}{\partial \tau^g} - s_{ivt} \sum_{v'=1}^{J} s_{iv't} \frac{\partial \phi_{iv't}}{\partial \tau^g} \\ &= \beta_i \gamma \rho m_{it} \; s_{ivt} \; (\frac{1}{mpg_v} - \sum_{v'=1}^{J} \frac{s_{iv't}}{mpg_{v'}}). \end{aligned}$$

Similarly, consumer i responds to the product tax by adjusting her choice probability such that

$$\begin{aligned}\frac{\partial s_{ivt}}{\partial \tau^{mpg}} &= s_{ivt}\frac{\partial \phi_{ivt}}{\partial \tau^{mpg}} - s_{ivt}\sum_{v'=1}^{J} s_{iv't}\frac{\partial \phi_{iv't}}{\partial \tau^{mpg}} \\ &= \beta_i \ s_{ivt} \ (\frac{1}{mpg_v} - \sum_{v'=1}^{J} \frac{s_{iv't}}{mpg_{v'}}).\end{aligned}$$

I compare expressions of consumer response to untangle the differentiated effects of the fuel tax and the product tax on consumer i's choice probability and hence the market shares of car models with different fuel economy ratings.

A.2 Data

After obtaining records of all non-commercial and non-diesel vehicles, I use the vehicle identification number (VIN) to construct a vehicle history sequence for every vehicle. I also use the combination of VIN and plate identifier to construct an owner history sequence for each vehicle. I mark a vehicle as a newly purchased one if its first vehicle history record is also the first owner history record, and the starting odometer reading of this record is smaller than 300 miles.[1]

I apply two criteria to flag low-quality observations. In MAVC, registration records are split where a mileage estimate begins or ends (Reardon et al., 2016). Registration periods without a corresponding mileage estimate are retained but assigned a “false” value for inspection matching. In addition, the temporal overlap between the mileage estimate and the registration record is compared to the length of the mileage estimate period as a measure of data reliability. A high value for the percentage of overlapping period suggests that the vehicle had the same owner and was garaged in the same location for a large portion of the mileage estimate period; a low value means that a substantial portion of the estimated mileage may have been driven while the vehicle was owned by another person or garaged in a different location. I flag a vehicle as a bad observation if there is a “false" value assigned or the percentage of overlapping period is smaller than 90% for any of its inspection records.

Dividing all new vehicles registered in 2011 which have no low-quality observation flags into two groups by registration time, I report the vehicle count by model year (MY) and by vehicle type in table A.1. Almost all passenger cars registered in the second half of 2011 are MY 2011/2012, while only 88% for new car registrations during the first half of calendar year 2011 are for MY 2011/2012.

I employ all MY 2011/2012 new passenger cars registered during the second half of 2011 without any low-quality observation flags as the demand sample for analysis.

[1] When not applying restrictions on the starting odometer readings, the observation pool constructed from matching two sequences includes brand new vehicles, vehicles released from commercial fleet, used vehicles came from other states, and those missed in previous census data. 200 miles is the other starting odometer reading tested for robustness. The results are similar.

Table A.1.
New Vehicle Registration Counts in 2011

	Jan - Jun 2011				Jul - Dec 2011			
Model Year	**Car**	**SUV**	**Truck**	**Van**	**Car**	**SUV**	**Truck**	**Van**
2010 and earlier	1,078	178	106	77	64	17	7	2
2011	7,788	4,882	1,120	549	5,119	3,488	1,091	405
2012	364	34	0	28	4,631	2,131	274	275
Total	9,230	5,094	1,226	654	9,814	5,636	1,372	682

From this demand sample, I select vehicles with exactly three consecutive inspection records since their first registration to form the mileage sample. The average number of days between two inspections is 379 for cars included in the mileage sample, and the standard deviation is 32 days.[2] Using the market-specific mileage sample, I construct the empirical distribution of expected annual vehicle miles traveled for new car buyers in each new car market.

In Massachusetts, thirteen Metropolitan Planning Organization (MPO) Regions showed in figure A.1 cover all 351 municipalities of the state. An MPO is a federally required regional transportation policy-making organization made of representatives from local government, regional transit operators, and state transportation agencies. Each MPO creates a fair and impartial setting for effective regional decision making in the metropolitan area to effectively engage communities and stakeholders.[3] Following the Massachusetts Travel Survey published by the Massachusetts Department of Transportation in 2012, I employ MPO regions to define the geographic new car mar-

[2]About 70% vehicles included in the demand sample show up in the mileage sample. The standard deviation of days between two inspections for vehicles with two or four consecutive inspection records is much larger.

[3]Refer to this website for more information about MPOs in Massachusetts:
https://www.mass.gov/service-details/regional-planning

ket. Tables A.2 and A.3 present vehicle counts by MPO and by registration month for demand sample and mileage sample accordingly.

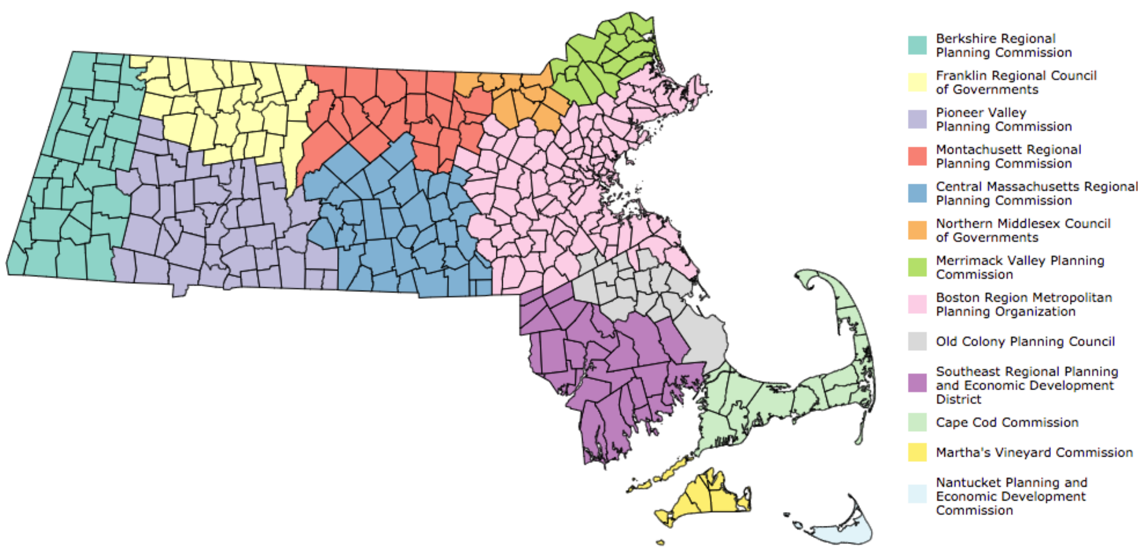

Figure A.1. An MPO is a federally required regional transportation policy-making organization made of representatives from local government, regional transit operators, and state transportation agencies. Each MPO creates a fair and impartial setting for effective regional decision making in the metropolitan area to effectively engage communities and stakeholders. Source: Massachusetts MPO Website Finder

Table A.2.
Vehicle Counts by Region and by Month, Demand Sample

MPO	**07/2011**	**08/2011**	**09/2011**	**10/2011**	**11/2011**	**12/2011**	**Total**
Berkshire	29	26	32	32	26	35	180
Franklin County	8	12	6	13	12	9	60
Pioneer Valley	100	115	115	95	106	114	645
Montachusett	48	51	52	46	39	45	281
Central Massachusetts	114	104	146	114	123	116	717
Northern Middlesex	65	76	59	72	67	63	402
Merrimack Valley	70	91	73	74	55	46	409
Metro Boston	735	958	890	805	842	874	5,104
Old Colony	75	74	96	73	73	70	461
Southeastern Massachusetts	103	122	113	104	95	96	633
Cape Cod	41	40	62	45	57	52	297
MarthaâĂŹs Vineyard	3	7	1	2	1	6	20
Nantucket	1	0	2	1	1	1	6
Total	**1,392**	**1,676**	**1,647**	**1,476**	**1,497**	**1,527**	**9,215**

Table A.3.
Vehicle Counts by Region and by Month, Mileage Sample

MPO	07/2011	08/2011	09/2011	10/2011	11/2011	12/2011	Total
Berkshire	16	15	20	22	16	25	114
Franklin County	8	9	4	12	7	7	47
Pioneer Valley	74	85	75	67	70	73	444
Montachusett	28	33	41	33	30	32	197
Central Massachusetts	81	77	93	85	80	77	493
Northern Middlesex	50	52	39	48	43	45	277
Merrimack Valley	50	72	57	57	31	30	297
Metro Boston	501	680	665	548	575	563	3,532
Old Colony	55	55	69	56	42	46	323
Southeastern Massachusetts	74	89	78	70	68	61	440
Cape Cod	25	25	45	32	38	32	197
MarthaâĂŹs Vineyard	2	6	1	1	0	3	13
Nantucket	0	0	2	0	1	0	3
Total	**964**	**1,198**	**1,189**	**1,031**	**1,001**	**994**	**6,377**

A.3 Estimation Details and Numerical Considerations

Algorithm 1 lists detailed steps of the demand model estimation.

Algorithm 1: Demand Model Estimation

```
1 Sample N individuals, each of which consists of a K + 1 dimensional set of
    shocks;
2 Set the starting value of mean utilities at a vector of guessed values δ̃, and the
    starting value for the vector of random coefficients at corresponding
    parameter estimates from a linear logit model θ̃;
3 while |δ̃ − δ| > ε, do
4     Use δ̃ and θ̃ to compute market shares, applying the set of individuals
        simulated in step 1 and equation 1.6;
5     Take difference between the logarithms of observed market shares and
        those calculated from previous step, mark it as Δ;
6     Update δ̃ to δ using Δ;
7 end
8 Compute error term vector ζ̃_vt(θ) using δ;
9 Construct the objective function of the GMM estimator, and search for θ to
    minimize it.
```

A.4 Revenue-equivalent Product Taxes

Once a new gasoline tax is applied to the original level τ, the change of total gas tax paid by consumer i driving car v becomes $\rho \frac{m_i}{mpg_v} \tau^g$ over the car's lifetime. The expected revenue from the new gasoline tax over all cars in the choice set per consumer in market t turns to

$$g_{it}(\zeta_{vt}, \tau, \tau^g; \alpha_i, \beta_i, \beta_i\gamma\rho) = \sum_{v=1}^{J} s_{ivt}(\zeta_{vt}; \alpha_i, \beta_i, \beta_i\gamma\rho)\ \rho\ \frac{m_{it}}{mpg_v}\ (\tau + \tau^g). \tag{A.1}$$

The revenue generated from the fuel tax in market t is

$$R_t = \int_\theta g_{it}(\zeta_{vt}, \tau, \tau^g; \alpha_i, \beta_i, \beta_i\gamma\rho)\ dF(\theta) I_t. \tag{A.2}$$

To compute a revenue-equivalent product tax, I derive an equation of expected revenue per consumer over all available car models for the product tax similar to equation A.1 and integrate that to the market-specific total revenue. Algorithm 2 explains this process in details.

Algorithm 2: Compute Tax Revenue

1. Set the new fuel tax and hence the new gas price;
2. Apply the same set of simulated individuals as used for the demand estimation;
3. Load the vector of residual error terms ζ_{vt} obtained from the demand estimation to hold the demand system constant;
4. Simulate new consumer choices given the change in gas prices, and compute the new choice probability for each car in the choice set of individual i;
5. Calculate the expected tax revenue over all cars per consumer using equation A.1, and take an average of the expected tax revenue over all consumers in market t;
6. Multiply the average expected tax revenue per consumer by the number of potential buyers in each market I_t to get the total expected tax revenue in market t as described in equation A.2;
7. Solve for a revenue-equivalent product tax by the market and obtain corresponding individual choice probability vectors.

A.5 Elastic Demand for Driving

Table A.4.
The Elasticity of Driving in Mileage Sample

	Two-way Fixed Effects with IV
ln(midgrade gas price)	-0.77*
	(0.35)
ln(unemployment rate)	0.23
	(0.14)
ln(national CCI)	4.74**
	(1.79)
ln(months between inspections)	0.05
	(0.04)
ln(summer proportion)	1.12***
	(0.33)
Constant	-17.97**
	(7.86)
Vehicle FE	Yes
Record-end-year FE	Yes
Record-end-month FE	Yes
R-square	0.04

Notes: The dependent variable is ln(average daily VMT). VMT stands for vehicle miles traveled. The number of observational units is 6,377. F-statistic for the first stage regression is 3.3×10^5. Values in parentheses are robust standard errors clustered at the MPO level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

B. APPENDIX: IDENTIFYING THE ELASTICITY OF DRIVING

B.1 Instrumental Variable for Massachusetts Motor Fuel Prices

Table B.1.
First-stage Regression Results, Two-way Fixed Effects Case

	2FE
ln(Brent Crude)	0.6508***
	(0.0009)
ln(vehicle age)	0.1817***
	(0.0188)
ln(MA unemployment)	0.0459***
	(0.0024)
ln(national CCI)	-0.1300***
	(0.0221)
ln(MA GDP)	-0.4880***
	(0.0138)
Constant	-2.5880***
	(0.0942)
Municipality FE	Yes
Quarter FE	Yes
F-statistic	2.3×10^5
R-square	0.99

Notes: The dependent variable is ln(fuel price per mile). Regressions are for N = 351 municipalities, n = 8,424 observations in the levels specifications. Values in parentheses are robust standard errors accommodate for heteroskedasticity and autocorrelation. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

VITA

VITA

Tingmingke Lu was born in 1987. He grew up in Changzhou, Jiangsu Province, China. Tingmingke Lu earned a B.S. in Mathematics and Applied Mathematics from East China Normal University in Shanghai, China. After working at Cisco System (China) R&D Center for two years, he attended Kent State University in Ohio where he earned his M.A. in Economics. Then he enrolled in the Doctoral Program in Economics at Purdue University where he earned a Ph.D. in Economics. In August 2019, he will join the Swedish University of Agricultural Sciences in Uppsala, Sweden, as a Post-doctoral Researcher.