Paris School of Economics
Paris, Spring 2019

Aggregate Response to Environmental Taxation

Effects of gasoline taxation on transportation demand and vehicle fleet composition using compositional data analysis

Liam Collins
Supervisor: Katheline Schubert

Master Thesis, Analysis and Policy in Economics

PARIS SCHOOL OF ECONOMICS

Abstract

Adjustments to levels of Pigovian taxation mechanisms hold significant potential to alter the structure of an economy. Focusing on vehicle miles travelled (VMT) and total vehicle registrations, I show how economies respond to such changes and discuss ways in which a policy-maker could incorporate these reactions into the decision-making process. My results show that U.S. consumers are much more averse to changes in a gasoline tax than tax-exclusive retail price changes when considering VMT. I estimate an elasticity of vehicle registrations with respect to the tax on gasoline of -0.208. Developing a static macroeconomic island model with exogenous VMT illustrates the effects of gasoline taxation on economic structures. In the spirit of this model, an econometric analysis using compositional data techniques on relative quantities of vehicle registration types demonstrates that gasoline taxes can substantially alter vehicle fleet compositions.

JEL classification – H23, O18, R48
Keywords – Pigovian tax, transportation demand, gasoline tax

Acknowledgements

I would like to thank my supervisor, Katheline Schubert, for advice while writing this thesis as well as her encouragement in pursuing the study of environmental economics. I would also like to thank Mouez Fodha for agreeing to be referee for my thesis defense. Special thanks to Dan Rust at the University of Bielefeld who prepared me for the last two years, along with countless peers whose support was vital to the validation of the Masters.

Contents

1 Introduction

Policymakers and environmentalists often face roadblocks when attempting to pass carbon tax legislation.[1] Recently, in France, proposed tax increases coinciding with rising fuel prices led to civil unrest across the country in the late months of 2018, further evidence of the political risks of proposing carbon taxation. Yet, despite these hurdles, a recent report by the UN Intergovernmental Panel on Climate Change (IPCC) warns that urgent changes to economic policies are necessary to limit the global temperature to 1.5° C above pre-industrial levels, the ambitious goal set by the Paris Agreement in 2016. Nordhaus (2017) uses the DICE (Dynamic Integrated Model of Climate and the Economy) model to obtain conservative estimates suggesting a social cost of carbon of roughly $37 per ton in 2019 dollars, despite no federal carbon tax in the United States. According to the US Department of Energy, the transportation sector is responsible for roughly 33.6% of all carbon dioxide emissions in the United States.[2] While gasoline seems like a natural candidate for a carbon tax, there is research implying that motor gasoline tax rates in the US are far below the optimal level (Parry and Small, 2005). As state-level gasoline tax rates climb upward and the executive branch considers an increase to the federal tax rate, it is of great importance that policymakers understand the aggregate response such changes have on the transportation economy.[3] Many researchers have analyzed the various channels through which individual consumers respond to gasoline tax increases, however a more robust analysis of an economy-wide response is due in the environmental economics literature.

In the US, gasoline is taxed at both federal and state levels, although in aggregate the post-tax prices faced by American drivers are still much lower than those faced by drivers in European countries. While the federal excise tax on gasoline has remained constant at 18.4 cents per gallon since 1993, there are significant differences in per-gallon tax levels across states. Much of the econometric literature exploits this variation to assess the role of gasoline taxes on consumer behavior in transportation markets. Davis and Killian (2011) find that a 10-cent increase in the gasoline tax would decrease total US emissions by about 0.5%. This estimate incorporates a short-run approach and does not capture changes to the fuel-efficiency of the aggregate vehicle fleet. In order to address long-run consumer adaptations, Klier and Linn (2010) look at a unique dataset of monthly vehicle sales from 1978 to 2007. They find that a $1 increase in gasoline price would increase the average fuel economy of vehicles sold by 0.8-1 miles per gallon.

Gasoline prices change due to a number of factors, including taxes and shocks to the world price of oil. The extent to which consumers incorporate these impacts in their purchasing behavior varies. A consumer may buy cheaper vehicles, drive less frequently, or even drive slower

[1] See for example, Washington State Initiative 1631, and the American Clean Energy and Security Act (HR 2454). The Washington measure would have enacted a fee of $15 on each metric ton of carbon emitted in the state of Washington, increasing by $2 per year, while the revenue would have funded investments in green technology. The measure failed 56% to 44%. HR 2454 would have established a carbon trading scheme similar to ETS in the EU. It passed the House but not was not brought to the Senate floor for discussion.

[2] US Department of Energy, *Transportation Energy Data Book*, 2017. cta.ornl.gov/data

[3] Jennifer Jacobs, Margaret Talev. Bloomberg News. May 2017. *Trump Open To Raising Gas Tax and Negotiating Tax Overhaul.* https://www.bloomberg.com/news/articles/2017-05-01/trump-open-to-raising-gas-tax-and-negotiating-tax-overhaul-plan

in an effort to curb transportation expenditures. To study the level of myopia in automobile consumers, Busse et al. (2013) look at the fuel-efficiency of new and used car purchases and pre-tax variation in gasoline price to shed light on the extent to which consumers incorporate future fuel costs into their decision-making process. Their analysis reveals that consumers incorporate rather fully these expected future burdens and buy more fuel-efficient vehicles to make up for the costs. On the other hand, Allcott and Wozny (2011) find that consumers equate a \$1 increase in car price today, with a \$0.72 increase in the discounted expected future costs of gasoline which implies a significant level of short-sightedness on behalf of consumers.

For any policy-maker who does wish to implement Pigovian taxation mechanisms, it is important to know the demand elasticity of the relevant polluting good. Dahl (1986) provides a gasoline demand survey, with an overview of various measures of transportation related elasticities. She examines own-price elasticities and income elasticities of gasoline quantity demanded, vehicle fleet fuel efficiency and vehicle miles traveled (VMT). She finds that estimates for own-price elasticity of VMT range from -0.1 to -0.5 in the short run and -0.5 to -0.6 in the long run when looking at annual data with no lagged variables. The estimates for income elasticity range from 0.15 to 0.47 in the short run and 0.54 to 0.66 in the long run. Relying on state-level annual data, the elasticities of VMT calculated in this paper concur with much of the previous research. Moreover, the analysis will go beyond a simple partial-equilibrium treatment of the demand for household-level VMT and attempt to shift the focus to aggregate responses to taxation at an economy-scale level.

Much of the existing literature looks at vehicle purchase data (both new and old), scrapping data, and gasoline consumption data using detailed household-level data.[4] While there is much insight to be gained from a close examination of census and survey data, this paper focuses on aggregate data to get a broader picture of how the economy at large responds to gasoline tax changes. In response to a change in the fuel tax, one would expect firms to change optimally their investment in transportation capital, local municipalities to adjust their public transport (bus) services in addition to changing consumer behavior with regard to new and used vehicle purchases. For the majority of my analysis, I restrict the attention to vehicle registration data, categorized by vehicle class. There is a richness to this approach in that it accounts for every single vehicle that is legally certified to be in the vehicle fleet for a given year. Research that focuses solely on new vehicle sales excludes valuable information by ignoring how used vehicles are treated in the economy. New car purchases, cross-state sales of used vehicles, scrapping rates, and moratoria of older vehicles are not captured in this data set. Of course, public registration data does not contain model-level information on vehicles and thus definitive conclusions regarding fuel-efficiency of the aggregate fleet are less attainable. Li, Timmins, and von Haefen (2009) are able to exploit a unique panel dataset with model-level registration observations for 20 MSA's (Metropolitan Statistical Areas). In their simulations, they find that a 10% increase in gasoline prices from 2005 levels will generate a 0.22% increase in fleet fuel economy in the short run and a 2.04% increase in the long run.

[4]See Hausman and Newey (1995); West and Williams (2003); Goldberg (1998); Bento et al. (2009) among others

The vehicle registration data that I use contains aggregate information on the total number of registrations in each state-year by class: automobiles, motorcycles, buses, and trucks. Registration is not costless; households must must pay a fee, and thus the decision to do so will require weighing the trade-offs between expected future costs. Studying how the total composition of vehicle fleet changes with respect to covariates allows for a broad study of the transportation sector of a given economy. This has important applications in the cross-section of urban planning and environmental economics.

The paper proceeds as follows. Section II provides an overview of gasoline taxation and prices in the US. Section III describes the data used in this paper. Section IV examines an econometric model of vehicle miles traveled. These results are used to measure and compare elasticities to the literature and build off of earlier works relying on similar data sources. Section V provides a theoretical discrete choice model of demand for transportation modes and discusses the role of government taxation policies in a world of exogenous VMT with heterogenous consumers. The model provides a theoretical justification for why an analysis of taxation effects should encompass the full range of economic activity. In Section VI, I perform an econometric analysis of the response in vehicle fleet composition to gasoline tax changes using statistical techniques from the compositional data literature. Section VII offers concluding remarks.

2 Background

The origins of a tax on gasoline in the United States date back to a state law passed by the Oregon state legislature in 1919. The state was in need of a new channel through which to raise revenue to fund road construction.[5] The policy spread across the country at the state-level and in 1932, the Congress passed the Revenue Act of 1932 which levied a federal tax on gasoline at the rate of one penny per gallon. The federal rate continued to rise throughout the 20th century, eventually reaching 18.4 cents per gallon in 1993 when President Clinton signed the Omnibus Budget Reconciliation Act. Revenues generated by this tax provide funding for the Highway Trust Fund which holds two accounts. The 'Highway Account' funds road construction projects, and the 'Mass Transit Account' supports investments in public transportation projects. The federal tax has remained at this level since 1993, thus decreasing by about 43% in real terms, as of January 2019. Revenues from the tax were $26 billion in fiscal year 2016.[6]

At the state level, every state and D.C. has its own laws establishing excise taxes, which vary significantly from year to year as state legislatures make adjustments to balance state budgets. The legislation that augments gasoline taxes generally takes several months to take effect. Occasionally, a schedule is put in place to raise rates over a longer period. Therefore, it is generally thought that the gasoline tax rates are uncorrelated with world oil prices. Davis and Killian (2011) use this intuition to develop instruments. They are critical of the econometric literature that uses gasoline prices to measure price-elasticity, without directly adjusting for the endogeneity of these variables, so they use as instrumental variables changes in taxes which are directly correlated with increases in at-the-pump prices, yet exogenous to changes in consumer demand or preferences. Doyle and Samphantharak (2007) on the other hand, exploit a tax moratorium in Indiana and Illinois from 2000 that was put in place as a direct response to high gasoline prices. Their analysis finds that only 70% of the reduction in price from the moratorium are passed on to consumers, yet prices increase by 80-100% of the tax when it is reinstated. This is a relatively rare natural experiment, possible only via a statute giving gubernatorial power to declare a state of emergency in response to high energy prices. The authors determine that given that short-run demand elasticities are known to be quite small, the less than full reduction in retail price suggests that the short-run supply response is inelastic as well.

Gasoline taxes are not the only mechanism used by governing bodies to address environmental concerns from the transportation sector. Federal agencies have enforced emissions standards that require manufacturers to meet certain criteria with regards to fuel efficiency. In 1975, as part of the Energy and Conservation Act, Congress established for the first time average fuel economy standards that each manufacturer's fleet must achieve. As a response to the 1973 oil embargo, the act requires all car manufacturers to achieve a sales-weighted level of fuel efficiency. The first CAFE (Corporate Average Fuel Economy) levels were applied only to passenger vehicles and mandated that average fuel efficiency reach 27.5 miles per gallon by 1985. Light-duty trucks

[5] Oregon Department of Transportation. https://www.oregon.gov/ODOT/FTG/Pages/About-Us.aspx

[6] U.S. Department of Transportation. Federal Highway Administration. https://www.fhwa.dot.gov/policyinformation/statistics/2016/fe10.cfm

were soon incorporated into the program and targets were raised slightly, although stagnated from 1990 until 2007, when Congress passed the Energy Independence and Security Act. The exact levels of the policy are set by the National Highway Traffic Safety Administration, within the Department of Transportation. Unique authority is given to the California Air Resource Board to set its own more stringent state standards, which have been adopted by 13 other states and D.C.[7]

Austin and Dinan (2004) compare the costs of using CAFE standards versus a gasoline tax to achieve a long-run 10% reduction in gasoline consumption. They find the necessary gasoline tax approximately 40% less costly than the necessary CAFE standards. Agras and Chapman (1999) study the two policies and measure increasing returns from using the two together in order to meet CO_2 emissions goals put forth by the Kyoto Protocol.[8] The Protocol required its signatories to reach 93% of their 1990 emissions levels by 2010. In order to attain this 93% goal within the transportation sector, they find that a CAFE policy would require an annual increase of 3.2 miles per gallon from 1999 until 2010. Under a gasoline tax policy, an annual increase in 10.5 cents per gallon is required. However, they find that, when used jointly, only 41% of each is needed to achieve the emissions goal.

The Highway Trust Fund is no longer solvent with respect to the revenues from gasoline taxation.[9] In order to raise revenue to further invest in infrastructure projects and with mounting environmental concerns, an increase in the federal gasoline tax is imminent. The next section presents in detail the data that will be used to investigate consumer response to gasoline taxation using measurements of VMT and tallies of vehicle registrations.

[7]However, in early 2018 the legality of this accommodation began to be scrutinized by the Trump administration. see https://www.nytimes.com/2019/04/10/climate/auto-emissions-cafe-rollback-trump.html

[8]While the US never ratified the treaty and therefore was never subject to its objectives, President Clinton signed the Protocol in 1998.

[9]More than $130 billion has been transferred from the general fund over the last decade. More info is available at: https://www.enotrans.org/article/ten-years-of-highway-trust-fund-bankruptcy-why-did-it-happen-and-what-have-we-learned/

3 Data

The data set used for this research contains state-level observations for all 50 states (as well as D.C.) and includes information on vehicle registration tallies, gasoline prices, gasoline excise taxes and state economic and social indicators for the years 2000 to 2016.[10] Vehicle registration data are categorized by vehicle class (auto, motorcycle, bus, or truck) and come from the Federal Highway Administration (FHWA)'s Highway Statistics Series. Automobile classification includes all passenger cars (sedans, coupes, and station wagons manufactured primarily for the purpose of carrying passengers); buses include all vehicles manufactured as traditional passenger-carrying buses with two axles and six tires or three or more axles; truck classification is broader than the colloquial definition of truck and tends to include most larger vehicles designed for transportation of material. A more in-depth description of classification methodology is available in the user guide provided online by the FHWA.[11] Owners of motor vehicles are required to register their vehicles with their states' Department of Motor Vehicles (or equivalent). The exact nature of this registration process differs across states and municipalities although all states require the payment of fees to do this. States then report registration data to the U.S. Department of Transportation which releases the aggregate data.

The Statistics Series also includes estimates of vehicle miles traveled by road type (urban or rural), total public road length, and gasoline and diesel tax rates. Vehicle miles travelled (VMT) are calculated and reported by the Federal Highway Administration and come from measurements taken by the Highway Performance Monitoring System and State administrations that measure changes in traffic flow from over 5000 locations across the country. The tax rates reported are per-gallon excise taxes and the exact dates of enactment are given. Ad-valorem taxes are ignored for this econometric procedure. For example, California has a gasoline sales tax rate of 2.25%, lowered from the typical rate of 7.25% for most commercial items, although most states do not have this enodogeneity issue.

Yearly population estimates are provided by the Federal Reserve Economic Data (FRED) database. State GDP estimates were collected from the Bureau of Economic Analysis (BEA). State-level average gas prices were gathered from the State Energy Data System which is organized by the Energy Information Administration. This highly developed data source provides state-year data on a range of energy quantities and prices. Tax rates, gasoline prices and GDP are presented in nominal terms.

Table 3.1 presents summary statistics for state-year observations of registration levels, gas prices, tax rates and other relevant descriptive variables. The significant variation in state taxation levels across years and states allows for a robust examination of the policy's effect on fluctuations in vehicle registrations. Over the 17-year period, the average number of year-to-year changes in tax rates for a state was 3.6. Half the states had more than 2 changes and 12 states experienced more than 5 year-to-year changes. Figue 3.1 presents the standard deviation in state

[10] Data for more recent years are not available. Furthermore, while registration data is available for years before 2000, several important covariates were not accessible so the analysis is restricted to these 17 years.

[11] https://www.fhwa.dot.gov/policyinformation/statistics/2016/userguide.cfm

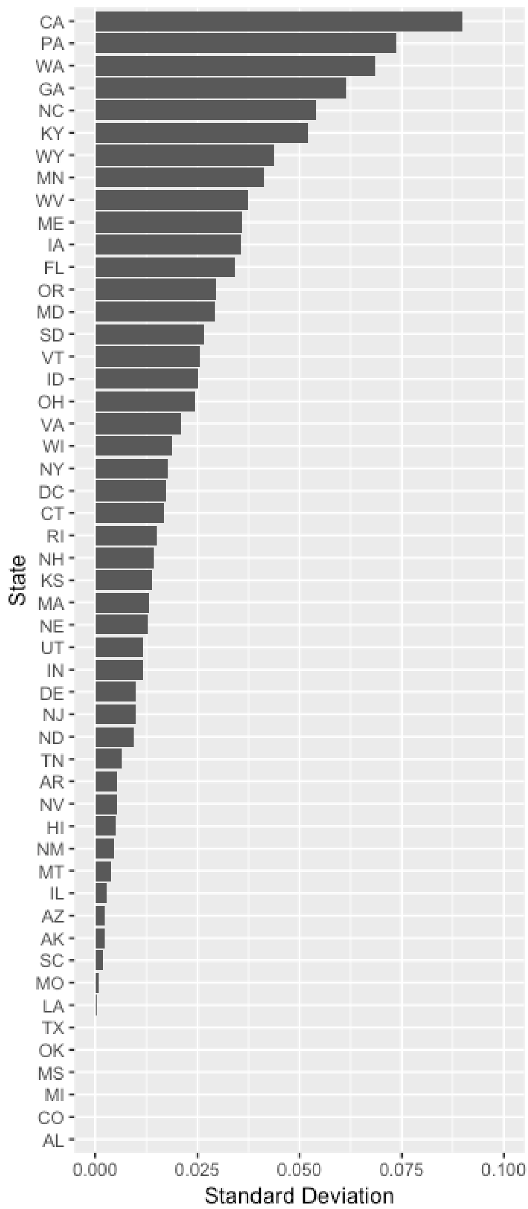

Figure 3.1: State variation in gasoline tax

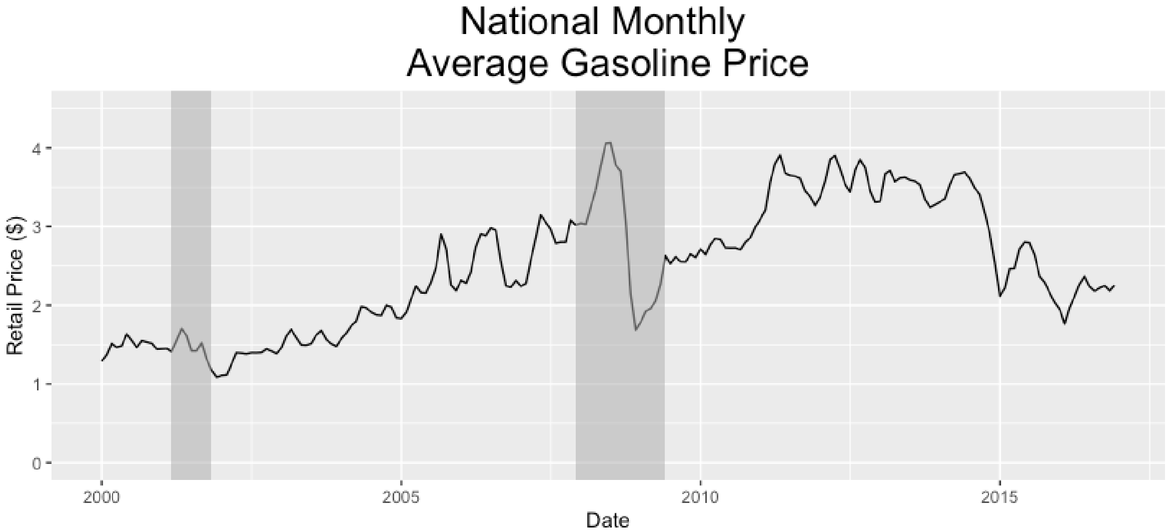

Figure 3.2: Gasoline Retail Price

tax rates for each state from 2000 to 2016. Six states (Texas, Oklahoma, Mississippi, Michigan, Colorado, and Alabama) have not had any changes to the tax rate in this period. California has the highest variation due primarily to sharp increases since 2011 designed to combat local pollution. Diesel is generally taxed at higher rates than gasoline, although diesel engine vehicles are much less common in the US than in European countries.[12] In 2000, diesel engine vehicles accounted for less than 1% of new retail car sales (DOT, Transportation Energy Databook). This number has risen through the 2000's but was still only 2.4% in 2014. For this reason, diesel taxation is not addressed explicitly in this paper's analysis.

Autos and trucks make up the vast majority of vehicle registrations, followed by motorcycles. Buses account for less than 1% of registrations in all states and years except for D.C. throughout the panel series, and Alaska in 2016. GDP fluctuates dramatically over states, as expected. D.C. is an outlier for GDP per capita due to its small size and the fact that it hosts most economic activity of the federal government. All other states have observations below 80,000 for this variable.

Also included in the data set are variables describing the political landscape of all states in each year, including the party in control of each state's legislature and the party in control of the gubernatorial office. Each state has a bicameral legislature (except Nebraska) and in Table 3.1, I report the percentage of state-year observations where Democrats or Republicans control both houses, as well as the the frequency of split control when each party controls one state house. This occurred in 18.5% of all observations. As is clear from the table, the Republican party has exercised dominant control of state-level politics throughout the panel series. In order to address the political structure in the econometric models, I develop a variable that measures the political lean of the state. It is defined simply as the sum of "entities" (legislative house or

[12]It is noteworthy to point out that diesel and gasoline tax levels are highly correlated: $\rho = 0.83$

governorship) controlled by Democrats. Thus, it serves as a discrete scale from 0 to 3, where 3 is complete Democratic control and 0 is complete Republican control. Washington D.C. has no legislative body but I assign it a score of 3 as it consistently votes Democratic in presidential elections and has had a Democratic delegate to the House of Representatives throughout the duration of the panel. Nebraska has a non-partisan, unicameral state house, but I have assigned it a score of 0 with similar reasoning.

Rural vehicle miles travelled is, naturally, 0 for D.C. in all years. Most states have higher levels of urban vehicle miles travelled than rural. In 2000, approximately 60% of all miles driven were in urban settings. By 2016, this figure had risen to 70%.

Gas price varies substantially throughout the time series. Figure 3.2 shows monthly averages of retail gasoline prices at the national level. The data for this chart comes form the U.S. Energy Information Administration and provides context for national trends stemming from fluctuations in supply and demand. Regression bands are added and explain two significant drops in gasoline price: the mild recession of 2001-2002, followed by the financial crisis of 2008 which accompanies a steep drop in gasoline prices. There is also substantial variation in gasoline prices across states. Hawaii had the highest tax-inclusive gas prices in 13 of the 17 years in the panel. From a supply-side perspective this makes sense due to the high fixed costs associated with delivering crude oil products to the remote island. In addition, a state law enacted in 2006, requires all motor gasoline sold in the state to be blended with 10% ethanol. Georgia consumers faced the lowest prices for 14 of the 17 years. Across years the median difference in average gas price between the most expensive and least expensive state was \$0.67, while in 2012, Hawaiians paid on average \$1.13 more per gallon of gas than Georgians, a sizeable discrepancy. This is also due in part to Georgia having the lowest gasoline tax in the sample from 2000 through 2014.

Table 3.1: Summary Statistics

Variable	N	Mean	St. Dev.	Min	Max
Auto	867	2,518,030	2,961,920	160,090	20,037,727
Trucks	867	2,227,672	2,372,134	39,164	14,511,913
Buses	867	15,993	18,355	1,092	101,845
Motorcycles	867	138,692	140,952	1,100	842,106
Gasoline Tax	867	0.219	0.061	0.075	0.505
Diesel Tax	867	0.225	0.071	0.075	0.642
GDP (billions USD)	867	279.48	351.7	17.2	2,665
Population	867	5,946,322	6,668,041	234,815	39,209,127
Land Area	867	69,348	84,881	61	571,951
Rural VMT (millions)	867	19,878	15,344	0	85,966
Urban VMT (millions)	867	38,238	47,623	1,834	287,120
Public Road Length	867	69,877	49,114	1,046	310,850
Gas Price (no tax)	867	2.16	0.8	0.89	4.27
GDP/capita	867	47,319	19,365	23,047	189,450
Political Structure:					
Legislature					
Democratic	39.6%				
Republican	42%				
Split	18.5%				
Governor					
Democratic	45.7%				
Republican	54.3%				

4 Empirical Analysis of Transportation

In order to understand how vehicle usage responds to taxation policies, I look at the effects of gasoline price tax rates on two measures of vehicle utilization: vehicles miles travelled and total vehicle registrations. While VMT modeling is fairly standard in the literature, registration data is a fairly neglected statistic for proxying transportation demand[13]. By studying both, I can compare how total transportation demand (VMT) reacts to changes in tax and price to how transportation capital (registrations) reacts to changes in tax and price.

4.1 Modeling Total Vehicle Registrations

Modeling of registrations will be done in two ways. In the first model, the post-tax price paid by consumers is the primary regressor. The linear equation for this model is:

$$\ln(R_{sy}) = \beta_0 + \beta_1 \ln(\tau_{sy} + p_{sy}) + \beta_2 \ln(GDP/capita_{sy}) + e_{sy}$$

In the second model, I isolate each of these components of the final price paid by consumers. The linear equation is then:

$$\ln(R_{sy}) = \beta_0 + \beta_1 \ln(\tau_{sy}) + \beta_2 \ln(p_{sy}) + \beta_3 \ln(GDP/capita_{sy}) + e_{sy}$$

where, in both equations, R_{sy} is the total quantity of vehicle registrations in state s in year y, τ_{sy} is the gasoline tax (inclusive of federal rate), and p_{sy} is the tax-exclusive retail price of gasoline. $GDP/capita_{sy}$ is GDP per capita and e_{sy} is the error term. In addition to these simple specifications I expand the models by adding controls and various covariates. First, I will discuss the results of the models that isolate price and tax.

The estimates from the first specification are given in column (1) of Table 4.1 along with two alternative specifications. It is evident that the simplest model (1) yields unsatisfying results. The coefficient on tax-exclusive price is positive while the coefficient on the tax level is insignificant. However, accounting for political structure, road length per square mile (which controls for the density of road infrastructure and thus captures some of the ability of households to utilize transportation technology), and year fixed effects offers coefficients that are much more tractable. Specification (2) implies an elasticity of vehicle registration with respect to gasoline tax of -0.39, while suggesting that consumers are much more elastic with respect to retail price, which takes a value of -4.757, an order of magnitude larger.

In specification (3), however, registration rates are found to be slightly more elastic with respect to gasoline taxes than to the tax-exclusive retail price. While a 10% increase in the tax rate would be expected to reduce the number of registered vehicles by 2%, a shock to the world price of oil that leads to a 10% increase in the pre-tax price, would be expected

[13]see Goldberg (1998) for modeling of VMT; Li et al. (2009) for modeling with vehicle registrations.

to reduce registrations by only 1.2%. All coefficients are statistically significant in this third model. The elasticity of GDP per capita is substantially higher than the first two specifications. In the simplest interpretation, one can recognize this value as the income elasticity of motor vehicles. Ignoring price effects, this means that as income increases uniformly across a set of households, some households who don't own any passenger vehicles will be pushed over a financial threshold that allows them to make the investment in a car, while some households further up the distribution will decide to purchase a second or even third vehicle. This coefficient takes a value of 1.214 and is highly statistically significant. Moreover, I can reject the null hypothesis that the true value is 1, and this supports the conclusion that household demand for vehicles is elastic with respect to income. However, the results are not very robust from this simplified thought experiment as the data concerned is better interpreted as output, not household income, and registration data consists of all vehicles, not only passenger cars. Taking into consideration the other aspects of vehicle demand that go beyond household interests in utility-maximization, such as firm-level transportation operations, one can interpret this value in a slightly broader fashion. Vehicles are not consumable goods in a direct way, but are used as (transportation) capital to obtain utility, produce output, or provide services. Therefore the elasticity of GDP per capita, being larger than 1, suggests that as economic output increases, the share of output resulting in transportation capital increases.

Table 4.2 shows the results from the same three specifications presented in Table 4.1 but now the variable concerning gasoline is the final at-the-pump price paid by consumers (sum of the tax-exclusive price and and the tax). Specification (3) shows a statistically significant price elasticity of -0.124. Li et al. (2009) calculated the elasticity of survival probabilities for various models across geographic regions for the year 2000. Having relied on a detailed model-level data set of vehicle registrations they were able to determine the probability that a vehicle would still be present in the local economy after a 1% increase in the tax rate. For instance, they found that a 1985 Chevy Impala (which has a fuel economy of 21.2) has survival elasticity of -0.1[14]. While such specific claims are impossible with the level of aggregation of my data, this provides an interesting benchmark that one can use to interpret the elasticity, -0.124.

Consider for a moment an economy with total registrations, R. The survival probability of all vehicles at time t is s. The wave of new cars is vR. A steady-state equilibrium is one in which $v = s$. Therefore, in expectation, R is constant. If the elasticity of s with respect to gasoline price is ϵ and the elasticity of v with respect to gasoline price is also ϵ then a 1% increase in gasoline price at time t would yield ϵ% fewer registrations at time $t + 1$. While this paints a quite simplified version of the market for vehicles, it depicts a nice illustration of how to view the elasticity of gasoline price in specification (3) with regards to the findings of Li et al. (2009) concerning survival probabilities.

[14] The average fuel economy for all passenger cars in 2000 was 21.9. Furthermore, the average age of a passenger car was 9.1 (DOT, Transportation Energy Book).

Table 4.1

	Dependent variable:		
	Vehicle Registrations (in logs)		
	(1)	(2)	(3)
log(tax)	−0.092	−0.392***	−0.208***
	(0.109)	(0.126)	(0.040)
log(tax-excl. price)	0.202**	−4.757***	−0.116***
	(0.093)	(0.622)	(0.017)
Political lean		−0.078**	−0.012*
		(0.035)	(0.006)
log(GDP/capita)	−0.639***	0.763***	1.214***
	(0.128)	(0.088)	(0.008)
log(Road length/mile^2)		0.077*	
		(0.045)	
State FE	No	No	Yes
Year FE	No	Yes	No
Time Trend	No	No	Yes
Observations	867	816	867
R^2	0.029	0.995	1.000
Adjusted R^2	0.025	0.995	1.000
Residual Std. Error	0.984	1.020	0.125
AIC	2437.981	2372.324	-1095.150
Note:	*p<0.1; **p<0.05; ***p<0.01		

4.2 Modeling Vehicle Miles Travelled

Using state-level aggregate data, Li, Linn and Muehlegger (2014) set up an estimable equation that isolates the tax-exclusive price of gasoline and the ratio of the tax to the price itself. The specification they use is given as:

$$\ln(q_{sy}) = \alpha \ln(p_{sy}) + \beta \ln(1 + \frac{\tau_{sy}}{p_{sy}}) + X_{sy}\Theta + \delta_s + \phi_y + e_{sy}$$

where q_{sy} is the relevant dependent quantity for state s in year y, p_{sy} is the tax-exclusive price level, τ_{sy} is the tax rate, X_{sy} is a vector of state-level characteristics, δ_s represent state dummies, ϕ_y represent time dummies and e_{sy} is an error term.

A critical feature of this modeling framework is its ability to test whether consumers respond identically to tax-related changes in gas prices as to equivalent changes stemming from supply-side effects. This can be demonstrated as follows. Assuming consumers bear the entire burden of the tax, we have that $\frac{dp}{d\tau} = 0$. This assumption is supported by the work described

Table 4.2

	Dependent variable:		
	Vehicle Registrations (in logs)		
	(1)	(2)	(3)
log(Gas Price)	0.220**	−3.572***	−0.124***
	(0.112)	(0.733)	(0.021)
log(GDP/capita)	−0.628***	−0.785***	1.255***
	(0.128)	(0.145)	(0.003)
log(Road length/miles^2)		0.211***	
		(0.041)	
Political lean		−0.046	−0.015**
		(0.032)	(0.006)
Observations	867	816	867
R^2	0.027	0.996	1.000
Adjusted R^2	0.025	0.996	1.000
Residual Std. Error	0.984	0.927	0.127
F Statistic	12.070***	10,597.200***	223,752.800***
Note:			*p<0.1; **p<0.05; ***p<0.01

in the previous section (Doyle and Samphantharak, 2007). It can be show that if consumers respond equally to gasoline-tax changes and tax-exclusive price changes, $\alpha = \beta$ and therefore, if consumers are more elastic to tax rates than to tax-exclusive price, β should be larger (in magnitude) than α. Taking the derivative of equation 4.1 with respect to the tax-exclusive gasoline price, p_{sy}, allows us to obtain the semi-elasticity, given by:

$$\frac{1}{p_{sy}}\Big(\alpha - \beta\frac{\tau}{p_{sy}+\tau}\Big)$$

Now, taking the derivative with respect to gasoline tax, and assuming $\frac{dp}{d\tau} = 0$, we obtain as the semi-elasticity:

$$\beta\Big(\frac{1}{p_{sy}+\tau}\Big)$$

These two results are equivalent if and only if $\alpha = \beta$. By using this specification we can use the variance-covariance matrix to test this linear hypothesis and examine whether consumers treat such price changes equivalently.

In their work, Li, Linn and Muhlegger (2014) use this equation to model gasoline consumption data, provided by the Federal Highway Administration. I apply their econometric formulation to measurements of total vehicle miles travelled. My analysis includes five different specifications. The first model is the simplest and contains just the estimates for α, β and a constant. In models (2) - (5), I include various combinations of covariates (logged GDP per capita, a political categorical variable, and logged public road length per square mile) in the design matrix, in addition to year and state fixed effects. Table 4.3 shows the results from these linear regressions. Specifications (1), (4), and (5) find significant results for both α and β. The estimated elasticity of GDP/capita is highly significant for all four specifications at the 1% level. Specifications (3) - (5) imply an elasticity between 0.15 and 0.5 which agrees with the results presented by Dahl (1986) discussed in Section I. Analyzing the effectiveness of fuel efficiency standards, such as CAFE, Mayo and Mathis (1988) find a long run elasticity of 0.3. As for the estimates of α and β, I obtain highly differentiated results across the first 3 specifications. Li et al. (2014) provide a table of 8 separate specifications when regressing on gasoline consumption, not VMT, but for their specification containing several covariates, year and state fixed effects, their results are -0.163 and -0.470, respectively, which is similar to the results I find for the full-fledged model in specification (5). Note that results analyzing gasoline consumption data, as is the case for Li et al. (2014), don't capture consumers' changing behavior with regards to demand for transportation explicitly. For instance, if an economy increases vehicular travel and this coincides with technology that improves fuel efficiency, gasoline consumption could remain constant, as these effects would to at least some cancel out each other. From a pure environmental perspective, an economist might not be concerned with changing VMT and would be only concerned with changing levels of gasoline consumption. However, more broadly, understanding VMT is important for other area of economic planning. Furthermore, for their model without year fixed effects, they obtain -0.078 and -0.331 for α and β, respectively, which are extremely similar to those obtained in specification (4).

In order to determine whether or not consumers treat differently a price change due to tax increase and an equivalent change in the tax-exclusive price, I perform a Wald-test on each specification to test the linear restriction that $\alpha = \beta$. P-values from the ensuing F-statistics are represented in the table. One can see that models (2) - (4) all reject the null hypothesis at a 5% significance level. I give specification (4) particular attention because it comfortably rejects this null hypothesis, and has significant coefficients that agree with findings standard in the literature. The results from this regression allow me to reject the null hypothesis that consumers are indifferent to a gasoline tax change and an equivalent change in tax-exclusive retail price, and I conclude that they are more averse to a change in gasoline tax than a change in tax-exclusive retail price, when considering vehicle miles travelled.

There are a variety of possible explanations for this result. Saliency from media coverage may be the best possible explanation. Media reports tend to cover such changes to policy as they are highly relevant to re-election prospects for incumbent politicians, while shocks to oil prices are common and are out of control of local policy-makers. As discussed earlier, President Macron faced significant backlash for the implementation of a tax hike on gasoline of 3.9 cents per litre following a 23% rise in world oil price. The media coverage of the ensuing protests brought widespread scrutiny on the situation not just in Europe but around the world. The coverage may have encouraged consumers to avoid driving or dissuaded households from taking vacations. The media coverage can serve as a mechanism of information dissemination. Another reason may be that tax increases today are indicative of tax increases in the future and part of a larger fiscal initiative while if oil prices are assumed to follow a random walk, in expectation gas price will stay the same forever.

Table 4.3: VMT Regression

	Dependent variable:				
	VMT (in logs)				
	(1)	(2)	(3)	(4)	(5)
log(tax-excl. price)	−0.627***	−0.356	−5.486***	−0.060***	−0.156***
	(0.242)	(0.235)	(0.689)	(0.020)	(0.045)
log(1 + tax ratio)	−2.969**	−4.800***	−0.959	−0.312***	−0.276**
	(1.295)	(1.259)	(1.377)	(0.107)	(0.125)
log(GDP/capita)		−1.198***	0.503***	0.238***	0.151***
		(0.135)	(0.088)	(0.019)	(0.026)
Covariates	No	Yes	Yes	Yes	Yes
State FE	No	No	No	Yes	Yes
Year FE	No	No	Yes	No	Yes
H: $\alpha = \beta$					
p-value	.055*	.00003***	.02**	.0052***	.25
Observations	867	816	816	816	816
R^2	0.008	0.168	0.991	1.000	1.000
Adjusted R^2	0.006	0.163	0.991	1.000	1.000
Residual Std. Error	1.017	0.933	0.992	0.041	0.039
F Statistic	3.42**	32.8***	4,159.3***	984,613***	821,912***

Note: *p<0.1; **p<0.05; ***p<0.01

5 Theoretical Model

5.1 Motivation

In contrast to modeling VMT directly as a function of economic characteristics, it would be interesting to construct a model in which VMT is not endogenous. The rest of this paper departs from the preliminary analysis modeling travel demand and instead examines how structural components of economies adapt to policy-measures when household demand for transportation is exogenously assigned.

Much of the existing literature studying vehicle choice decisions relies purely on household data and analyzes the decision-making process within a partial equilibrium framework. However, minimal consideration has been given in these models to the ways in which firms use transportation capital and governments provide transportation services in explicit relation to gasoline taxation. In this section, I develop a discrete choice model in which agents choose a transportation method (motorcycle, auto, or bus) to optimize individual utility. McFadden (1974) established the conditional logit model which became the central workhorse for modeling discrete choice behavior, by building upon random utility models in classical consumer theory. The main idea was that, by employing data on individual decision-making in which agents make a choice j where j is an element of the choice set B an econometrician can map individual i's characteristics and the attributes of alternatives $j \in B$ through a utility function. Utility can then be inferred by observing an individual's decisions, and an appropriate probabilistic structure can be obtained. Using a semi-parametric structure for utility and making distributional assumptions for the error term, McFadden (1974) shows the conditional choice probabilities have a multinomial logit form. This specification allowed for a highly tractable approach to discrete choice analysis, and has been used consistently in models of vehicle choice.

While this approach is well-designed for modeling empirical choice behavior, I wish to work with a theoretical model in which individual choices are not isolated phenomena, but instead create externalities. Parry and Small (2005) construct an intricate model of aggregate demand for vehicle miles travelled (VMT) using a representative agent to analyze an optimal level of gasoline taxation. The model presented in this section draws on theirs, although in my approach, VMT is exogenous, and households make decisions over consumption, labor, leisure and their vehicle choice to optimize utility subject to appropriate budget constraints. Models incorporating endogenous VMT, often assume a utility function increasing and concave in VMT.[15] The general rational is that households can travel for recreational purposes which is utility-enhancing. On the other hand, when travel is assumed to yield disutility, for example when households commute to work, it is often assumed that primary response mechanisms include residential or employment relocation.[16] This is often not the case as households face can strong geographic constraints. Furthermore, in 2009, 28% of all household vehicle miles travelled were for work commutes, so

[15]see Golob, Beckmann, and Zahavi, (1980)
[16]see McFadden (1975)

the notion that household utility is monotonically increasing in VMT is highly questionable.[17] This motivates establishing a model that exogenizes vehicle miles travelled across households.

5.2 The Model

The model consists of an island economy with N heterogeneous households, a firm that uses labor supplied by the island's households and capital from a world market, a benevolent environmentally-conscious government that controls taxation policy and a local council that makes transportation planning decisions. Households' heterogeneity is two-dimensional. Each household differs in the distance that its representative agent must travel to get to work, which is exogenously assigned, and its skill level which corresponds to individualized wage levels. One can picture this heterogeneity in the framework of a Hotelling model in a Cartesian coordinate system in which households are denoted $i \in \mathbb{R}^2$ (Hotelling, 1929). Location on the x-axis is the necessary VMT (distance to work, for instance) and position on the y-axis is wage level. Agents have three options when it comes to transportation. They can choose between riding a motorcycle, driving an automobile or taking the bus, and thus they must make a discrete decision for vehicle $v \in \{m, a, b\}$. If an individual chooses automobiles, or motorcycles, he/she must pay a fixed cost registration fee, and pay variable costs in the way of a per-unit price of fuel, q_F and a per-unit excise tax, τ_F. Fixed costs are denoted Z_v, for vehicle choice $v \in \{a, m\}$, where $Z_a > Z_m$. Per-unit cost of travel for autos and motorcycles, respectively, are :

$$c_a = \frac{(\tau_F + q_F)}{FE_a}$$
$$c_m = \frac{(\tau_F + q_F)}{FE_m}$$

FE_a and FE_m denote the fuel efficiency (in miles per gallon of fuel) of automobiles and motorcycles, respectively, and $FE_a < FE_m$. If the individual chooses to use the bus, he/she must pay a fixed cost per ride, c_b. It follows therefore that the individuals' total travel expenditure is:

$$e_i = \begin{cases} c_m m_i + Z_m & \text{if } v = m \\ c_a m_i + Z_a & \text{if } v = a \\ c_b, & \text{if } v = b \end{cases}$$

Note that bus expenditure is not a function of VMT but fixed. This is often the case in most urban bus systems where passengers pay a fixed fee to ride the bus regardless of how long they stay on. Prices for transfers in urban systems and prices for long-haul journeys are generally marginally diminishing as distance or number of legs increases so modeling bus expenditures as linear in distance does not seem necessary. A fully parameterized treatment of the model would consider perhaps that with x working days, regardless of distance to travel, travel expenditure for bus rider i is $c_b x$ if $l_i > 0$ and 0 otherwise, where l is labor. Fixed costs are assumed to

[17] US Department of Energy, Transportation Energy Data Book, 2017. cta.ornl.gov/data

cover registration fees, title fees, rents, insurance and depreciation of vehicle valuation. The burden of paying the price for a car is implicitly incorporated into the fixed costs. A household unconstrained with respect to competitive capital credit markets is indifferent between buying now and selling later, or renting, so the fixed costs incorporate this financial burden by way of accommodating depreciation and rents. Capital markets on the island are not included in this model but an extension that considers this would be interesting.[18]

Utility for household i defined by $(m_i, w_i) \in \mathbb{R}^2$ that chooses transportation method v is given by:

$$U_i = u(c_i, n_i) - \phi(P) - \delta_A(R) \;\mid\; v_i$$

where c is the numeraire representing all non-transport related consumption, P is the aggregate level of pollution, and A represents severity-adjusted traffic accidents. Direct utility $u(.)$ is quasi-concave, whereas $\phi(.)$ and $\delta_A(.)$ are weakly-convex. The household faces two constraints. The budget constraint is:

$$c_i + e_i(v_i) = (1 - \tau_L)w_i l_i$$

where τ_L is the labor tax rate, w_i is the wage level and l_i is the total amount of hours worked. In addition, the household faces a time constraint which implicitly captures utility lost from needing to commute to work. This constraint is:

$$l_i + n_i + t_i(v_i) = \bar{l}$$

where t_i is the time spent traveling, and n_i is leisure. The total time endowment, $\bar{l}$, is the same for all households. Time spent traveling is an increasing and convex function in the total amount of vehicle registrations which represents the total level of congestion, determined by:

$$t_i = \gamma_v m_i R \log(R)$$

where R is the total number of vehicles registered and is equal to the sum of registrations for each vehicle type (i.e. $R = R_M + R_A + R_T + R_B$). R_T is the quantity of truck registrations and is discussed below. m_i is as defined above, and γ_v is a scaling parameter for each transportation method chosen by individual i where $\gamma_m < \gamma_a < \gamma_b$, representing the fact that motorcycles are faster than automobiles which are faster than buses. Motorcyclists are able to move through traffic very efficiently and less constrained when it comes to finding parking. Buses are the slowest option in that they have to stop continually to pick up and drop off other passengers.

[18]This approach would take the form of a dynamic capital-accumulation model with households making decisions over leisure, labor, consumption, savings, and transportation choice, with a temporal discount factor, β. If travel for leisure is also added, one could feasibly arrive at interesting solutions in which those with large capital holdings do not have to work and can travel for pleasure, polluting the island, while the poor must continue to work. This full-fledged model is beyond the scope of what is necessary for this paper, but interesting to consider nonetheless.

The pollution level P is defined as the emitted quantity of carbon dioxide in tons. It is a linear function of total gasoline usage, and is thus indirectly a function of vehicle registrations and VMT appropriately weighted in correspondence to fuel efficiency levels.[19]

The severity-adjusted traffic accident component is also determined by total vehicle registration levels and the transportation method chosen by the household.

$$\delta_A(R) = \kappa_v R^{\xi}$$

where $v \in \{m, a, b\}$ is such that $\kappa_m > \kappa_b > \kappa_a$, and ξ is positive. The more vehicles registered and thus on the road, the more dangerous is travel. The scaling parameter κ_v captures the increasing risk for motorcyclists, while buses are considered the safest. Partly due to their slower speed and the fact that they are driven be a trained professional.

There is a firm sector of the economy that produces output following a Cobb-Douglas technology. The firms use labor, transportation capital (trucks) and non-transportation capital. Trucks are "produced" from the aggregate capital stock and the remainder is used as non-transportation capital. This production process can be written:

$$Y = F(\tilde{K}, L, T) = \Lambda \tilde{K}^{\alpha} L^{\beta} T^{\gamma}$$

$$T = f(K - \tilde{K})$$

where K is total capital, $\tilde{K}$ is non-transportation capital, L is aggregate labor, and T is aggregate transportation capital (trucks). The parameters α, β, and γ are all positive. The production function exhibits constant returns to scale (i.e. $\alpha + \beta + \gamma = 1$). Production is done solely to provide for the consumption needs of the local economy. Capital used by the firm sector is assumed to be available from the global market while labor is provided by the local economy and is equal to $\sum_{i=1}^{N} l_i$ over all N households. Labor payments from the firm are then $\sum_{i=1}^{N} l_i w_i$. The rental wage of capital is r. Trucks used within the production process are used to ship intermediate and final goods around the local economy. Because trucks are necessary to provide consumable goods to the households, consumption increases pollution through the channel of firm transportation mechanisms. The low fuel economy of trucks is particularly damaging to the environment through its contributions to P and in this way, the transportation process reveals that all consumption is dirty.

The government collects revenue in the form of the environmental tax on gasoline and the labor tax. It uses this revenue to provide the bus service as public transportation. The resulting government budget constraint is therefore:

[19] The EPA estimates that roughly 112.5 gallons of gasoline combusted in automobile vehicles emits one ton of CO_2. Diesel engines emit CO_2 at a slightly higher rate. see Environmental Protection Agency, Greenhouse Gas Emissions form a Typical Passenger Vehicle. May 2014. Office of Transportation and Air Quality (EPA-420-F-14-040).

$$\tau_F F + \tau_L \sum_{i=1}^{N} l_i w_i = G(B)$$

where $G(B)$ is the minimum government expenditure needed to supply B buses. F is the sum of all gallons of gasoline used in the economy. Viewed anther way, the amount buses available in the economy is,

$$B = G^{-1}(\tau_F F + \tau_L \sum_{i=1}^{N} l_i w_i)$$

The local council plays no role in the direct provision of services (transportation or otherwise) but makes decisions that improve the flow of economic activity on the island. In particular, it can design design a transportation-based agenda to optimize roadways the movement of vehicles. The council's agenda may include setting up bus-only lanes, directing guidelines for installing parking structures near the firm, and arranging the installation of gas stations. The local council needs to know the relative quantities of vehicles on the road in order to know how to optimally organize the island. In the next section, I will show how an equilibrium in this model is characterized.

5.3 Equilibrium and Discussion

Solving for a welfare-maximizing equilibrium of the economy describe in section 5.1 is a highly sophisticated discrete optimization problem. In this section, I will discuss the individual's maximization program, the firm's role in producing for aggregate demand under profit-maximization, and the policy mechanisms available to the benevolent government in order to increase aggregate social welfare.

Let us examine household i defined by (w_i, m_i) who observes all other households' transportation decisions. Conditional on having chosen $v \in \{a, b, m\}$ their optimization problem is:

$$\begin{aligned} \max_{c,l,n} \quad & u(c, n) - \phi(P) - \delta_A(v|R_{-i}) \\ \text{subject to} \quad & c + e(v) = (1 - \tau_L) w_i l \\ & l + n + t(v|R) = \bar{l} \end{aligned}$$

where $e(v)$ is defined as before. Setting up this problem as a Lagrangian and using substitution for the first constraint yields:

$$\mathcal{L} = u\big((1 - \tau_L) w_i l - e(v),\ n\big) - \phi(P) - \delta_A(v|R) + \lambda(\bar{l} - l - n - t(v|R))$$

where λ is the Lagrangian multiplier for the time constraint. We ignore the individual's direct effect on P and A and therefore the first-order conditions are:

$$
\begin{aligned}
&[1] \quad (1-\tau_L)w_i u_c - \lambda = 0 \\
&[2] \quad u_n - \lambda = 0 \\
&[3] \quad n + l = \bar{l} - t(v|R)
\end{aligned}
$$

Thus for any choice v, individual characteristics (w_i, m_i) and observation of vehicle choices for the $N-1$ other households, conditions [1] - [3] offer a well-defined decision rule.[20] Let the solution to these FOC's be denoted $(c^{opt}, n^{opt}, l^{opt})$ and let R_{-i} be the total observed registrations resulting from the vehicle decisions of the $N-1$ households. Individual i then chooses $v \in \{a, b, m\}$ that maximizes:

$$U(c^{opt}(v), n^{opt}(v), l^{opt}(v) \ \mid \ R_{-i}, m_i, w_i)$$

Given the choice decisions over all consumers, one obtains four N-dimensional vectors for consumption, labor, leisure, and transportation choices: $\{c_i\}_{i=1}^N, \{l_i\}_{i=1}^N, \{n_i\}_{i=1}^N, \{v_i\}_{i=1}^N$. The production side of the economy uses aggregate labor, $L = \sum_{i=1}^N l_i$ as the labor input to deliver output $Y \geq \sum_{i=1}^N c_i$. Outside capital, K, is divided between transportation capital, T, and non-transportation capital, $\tilde{K}$, as defined before. It follows that the transportation production is:

$$T = \left(\frac{Y}{\Lambda \tilde{K}^\alpha L^\beta}\right)^{(1/\gamma)}$$

As T has already been defined to be a function $f(K - \tilde{K})$, total capital used by the firm can be written:

$$K = \tilde{K} + f^{-1}\left[\left(\frac{Y}{\Lambda \tilde{K}^\alpha L^\beta}\right)^{(1/\gamma)}\right]$$

The firm's profit maximization is:

$$
\begin{aligned}
\max_{L,K,\tilde{K}} \quad & \Lambda \tilde{K}^\alpha L^\beta T^\gamma - rK - \sum_{i=1}^N w_i l_i \\
\text{subject to} \quad & T = f(K - \tilde{K})
\end{aligned}
$$

Despite, differing wage levels, all labor is equivalent within the production function. Optimization must account for this aspect of the model and the Lagrangian is therefore:

[20]Note: R and N are considered large enough that household i does not take into consideration in its own impact on P or severity-adjusted accident risk through F and R.

$$\mathcal{L} = \Lambda\tilde{K}^{\alpha}L^{\beta}T^{\gamma} - rK - \sum_{i=1}^{N} w_i l_i + \lambda(T - f(K - \tilde{(}K)) + \mu(L - \sum_{i=1}^{N} l_i)$$

where λ is the Lagrangian multiplier for the constraint for the production of transportation capital, and μ is the Lagrangian multiplier that labor in the production process is the sum of all labor. As wages vary, direct optimization is not possible without this constraint.

Government revenue is $\tau_L \sum_{i=1}^{N} w_i l_i + \tau_F F$ where and F is the total amount gallons consumed by automobiles, motorcycles, buses and trucks. Suppose σ is the number of individuals that a single bus can serve. Therefore the total number of buses required in the economy is $B = \frac{1}{\sigma}\sum_{i=1}^{N} \mathbb{1}_{\{v_i=b\}}$. The government provision feasibility constraint is:

$$G^{-1}(\tau_L \sum_{i=1}^{N} w_i l_i + \tau_F F) = \frac{1}{\sigma}\sum_{i=1}^{N} \mathbb{1}_{\{v_i=b\}}$$

Total pollution stems from the gallons of gasoline used by all the transportation channels described under the above conditions. We can write this:

$$F = \frac{\sum_{i=1}^{N} \mathbb{1}_{\{v_i=m\}} m_i}{FE_m} + \frac{\sum_{i=1}^{N} \mathbb{1}_{\{v_i=a\}} m_i}{FE_a} + \frac{1/\sigma \sum_{i=1}^{N} \mathbb{1}_{\{v_i=b\}} m_i}{FE_b} + \frac{h(T,Y)}{FE_T}$$

where $h(.)$ is the number of miles traveled by all trucks T which are needed to deliver $Y = C = \sum_{i=1}^{N} c_i$ consumable goods. It is increasing in both arguments. Fuel efficiency, as before, is defined as miles per gallon.

The total amount of vehicle registrations in the economy is $R = R_M + R_A + R_T + R_B$. The behavior of the consumers and the firm respond to two mechanisms in the government's policy arsenal: the gasoline tax, τ_F and the labor tax rate, τ_L. In this simplified model, government provides no other services. Aggregate social welfare from the perspective of the benevolent government is:

$$W = \sum_{i=1}^{N} U_i$$

Consumer actions create direct externalities that can be summed up in four ways.
1) Vehicle choice contributes directly to pollution via the fuel economy of this choice and VMT.
2) Increased consumption increases pollution because of the dirty production process (truck transportation). 3) Increased leisure means less tax revenue that is able to fund provision of bus services, the cleaner transportation channel. 4) An additional vehicle registration increases the traffic, and thus traffic time, decreasing leisure and/or labor of others. The government will try to maximize aggregate social welfare and is able to use the taxation policies at its disposal to do this.

From the formation of the above equilibrium registration quantities will respond directly to changes in tax rates. The local council monitors closely the changing ratios of vehicle registrations in response to the governments taxation mechanisms. Therefore, one can write:

$$R(\tau_L, \tau_F) = R_M(\tau_L, \tau_F) + R_A(\tau_L, \tau_F) + R_T(\tau_L, \tau_F) + R_B(\tau_L, \tau_F)$$

The ratios of these vehicle registration totals are of key interest in Section 6, in the spirit of the local council's decision framework. I am curious to analyze the development and response of the variables $(r_M, r_A, r_T, r_B) \in \mathbb{S}^4$ to marginal adjustments in policy, where $r_v = R_v(\tau_L, \tau_F)/R(\tau_L, \tau_F)$. The relevant policy mechanism that will be isolated and studied in an econometric framework is the gasoline tax rate.

5.4 External Validity

The original motivation for such a model is born from a skepticism of models that express VMT as purely enodengous. Furthermore, some models treat utility as increasing in VMT while others treat utility as decreasing in VMT. Neither assumption holds strictly in any feasible economic application, and therefore I feel a model that incorporates this variable as exogenous and unrelated explicitly to utility levels is warranted. Nonetheless, there are certainly many aspects of the model described above that limit its ability to accurately describe real-world economies.

First, the level to which households can substitute the three transportation methods is over-stated in the model. There are many factors that drive consumers to prefer motorcycles over automobiles beyond budgetary, temporal, and safety concerns. These include the number of individuals in the household and a range of unobservables that stimulate natural preferences. Households with passenger cars can carpool.[21] Motorcycles, on the other hand are clearly more limited in terms of capacity. Similarly, while heterogeneity of wages is captured in the model, work itself in real economies is heterogeneous and can require different means of transport.

Secondly, each of the three transportation methods does not represent a homogeneous transportation channel; automobiles vary greatly. A more robust model would incorporate a greater number of categorical segments and utilize stochastic elements within the household utility function to capture unobservable characteristics.

Thirdly, in the model, trucks serve uniquely as inputs of the firm production process signifying dirty consumption. In reality, and in the data to be analyzed in Section 6, categorization of trucks involves a much wider range of vehicles types than big-rig trucks that are used for transportation of consumption goods. Many businesses and local service providers use smaller trucks, and even some households whose recreational preferences require larger vehicles. Similarly, the model uses bus registrations to proxy for all forms of public transportation. Commuters use light rail, heavy rail, and commuter rails among other public transportation systems tin

[21] However, less than 10% of commuters carpooled to work in 2013, despite 86% of Americans using personal vehicles to get to work. US Department of Energy, *Transportation Energy Data Book*, 2017. cta.ornl.gov/data

addition to buses. Extending the model, I would like to include ridership of other public systems in addition to registration data to establish a more complete model of transportation demand response to gasoline tax changes. Despite the limitations mentioned above, a model that expands the effects of gasoline taxation beyond household decision-making provides new insight into the consequences of gasoline taxation.

6 Compositional Analysis

In section 4.1, I analyzed the response of aggregate vehicle fleet measured by total registrations to a change in gasoline tax rates by regressing the quantity of total vehicle registrations. In this section, as a means of applying the theoretical notions established in the discrete choice model of Section 5, I study the empirical data from a compositional perspective. In lieu of focusing on the total quantities of each vehicle type present in the registration data, the econometric approach here concerns their relative quantities. Armed with a better understanding of how vehicle fleet composition responds to changing gasoline tax and price levels, a policy-maker can improve how he/she approaches important economic planning questions. For instance, decisions around installations of bus-only lanes, high-occupancy vehicle (HOV) lanes, or high-occupancy toll (HOT) lanes can rely on an analysis of relative quantities across vehicle type not just total quantities or characteristics within vehicle type. In the following section, I provide a technical overview of the relevance of compositional data analysis in the statistics literature, and in section 6.2, I use these techniques to estimate parameters that will then provide inference for vehicle fleet composition changes across states.

6.1 Overview of CDA

Compositional data are quantitative data whose components consist of different categorical types that sum to some whole. Most often, these types of data present themselves as percentages, or ratios. A geologist, for instance, may wish to study the chemical or mineral compositions of a rock sample. An economist may wish to study how market shares of an oligopolistic industry change over time. An environmental scientist may want to study the changing levels of anthropogenic carbon dioxide in parts per million (ppm). If this type of relationship between components is of econometric interest, the dependent variable consists of a D-dimensional vector in the simplex, $\mathbb{S}^D$. John Aitchison has contributed immensely to the literature on compositional data analysis, and warns of the issues that arise from applying to such sample spaces techniques intended for real Euclidean space. Aitchison (1986) points to the fact that as raw quantities are irrelevant in compositional data analysis, it is the ratios of components on which statistical tools must be applied. He notes that this warning pre-dates his own research, citing Pearson who stressed the need for statisticians to "[be weary] of attempts to interpret correlations between ratios whose numerators and denominators contain common parts" (Pearson, 1897). Standard regression techniques on unconstrained space leads to these questionable interpretations. Aitchison refers to, specifically, the negative bias problem. Let a composition be $x = (x_1, \ldots, x_D)$ where $x_1 + \ldots + x_D = 1$. Due to the restriction that

$$Cov[x_1, x_1 + ... + x_D] = 0$$

it follows that

$$Cov(x_1, x_2) + \ldots + Cov(x_1, x_D) = -Var(x_1)$$

Thus, it is required that at least one term on the left hand side be negative, unless it is the trivial case that x_1 is a constant. This restriction is the reason why interpretability of parameters estimated by traditional statisitcal processes is greatly mitigated. Furthermore, researchers who are sensitive to this issue realize that there is no exact relationship between $Var(x_1/x_2)$ and $Var(x_2/x_1)$. But as ratios are the relevant statistical measures in compositional data analysis, this is unsatisfactory. Log-ratios, however, behave nicely:

$$Var[\log(x_1/x_2)] = Var[\log(x_2/x_1)]$$

Log-transformations therefore are called upon in much of the compositional data literature. Log-transformations from the simplex $\mathbb{S}^D$ to $\mathbb{R}^{D-1}$ allow the application of familiar statistical tools, as the space becomes unconstrained.

A common transformation presented by Aitchison (1986) is the additive log-ratio transformation. For a compositional vector $u \in \mathbb{S}^D$, an element j of the vector is selected and the transformation is:

$$alr(u) = (\log(\frac{u_1}{u_j}), ..., \log(\frac{u_{j-1}}{u_j}),\ \ \log(\frac{u_{j+1}}{u_j}), ..., \log(\frac{u_D}{u_j}))$$

Aitchison (1986) offers a proof demonstrating the irrelevance of the choice of j. Another transformation mapping objects from the simplex to real space, but that does not alter the dimension, is the centered log-ratio transformation. This transformation invovles dividing each component of the composition by the vector's geometric mean and then taking logarithms. For a compositional vector $u \in \mathbb{S}^D$:

$$clr(u) = (\log \frac{u_1}{g(u)}, ..., \log \frac{u_D}{g(u)})$$

where $g(u) = (\prod_{i=1}^{D} x_i)^{1/D}$.

Egozcue, Pawlowsky-Glahn, Mateu-Figueras, and Barcelo-Vidal (2003) extend this earlier work of Aitchison and propose using isometric transformations, the motivation being that angles and distances measured in the simplex can be associated to their counterparts in real space, under transformations that follow isometric properties. The isometric logarithmic transformation developed by the authors is as follows. The function ilr maps a vector $u \in \mathbb{S}^D$ to a vector $w \in \mathbb{R}^{D-1}$ where:

$$w_k = \sqrt{\frac{k}{k+1}} \log \frac{\sqrt[k]{\prod_{j=1}^{k} u_j}}{u_{k+1}}, \qquad 1 \leq k \leq D-1$$

It is then possible to apply standard linear regression techniques on the transformed data. One can obtain estimates for coefficients by employing multivariate least-squares regression. However, interpretations of the estimated parameters are not straightforward. The parameters themselves exist in the transformed space and must be handled carefully. Morais, Thomas-Agnan, and Simioni (2017) use this technique to study market shares of European automobile manufacturers. Their primary interest is to quantify the effect of the shares of advertising going to each medium (TV, radio, press, outdoor, internet and cinema) for each firm in an oligopolistic industry. They then measure the impact of these investment compositions on the firms' market shares over time.[22]

Incorporating this methodology, I construct a linear equation to be estimated by least-squares in the following way. Let $y_{st} \in \mathbb{S}^4$ be a compositional dependent variable vector representing the shares of vehicle registrations in year t and state s for motorcycles, automobiles, trucks, and buses. Let y^*_{st} be the data transformed by the isometric transformation specified above. That is, $y^*_{st} = ilr(y_{st}) \in \mathbb{R}^3$. Note that the choice of the "dropped" component is irrelevant. In other words, the order of the components within the vector have no bearing on the transformed estimated coefficients under OLS. Given the specified order of the components selected here, the D^{th} component represents buses. Furthermore, let j indicate the j^{th} component, for $j \in \{1, 2, 3\}$. Then the linear equations to be estimated are:

$$y^*_{jst} = a^*_j + \mathbf{c_j^*} \mathbf{Z_{st}} + {\epsilon_{jst}}^*$$

where y^*_{jst} is a the j^{th} component of the 3-dimensional transformed vector. a^*_j is a scalar and $\mathbf{c_j}^*$ is a vector of k parameters. Z_{st} is a vector of variables pertaining to state s in year t, and ϵ^*_{jst} is the error term. An asterisk denotes that the corresponding entity is to be interpreted only in the transformed space. Morais, Thomas-Agnan, and Simioni (2017) show that interpretation of coefficients of classical regressors on compositional dependent variables is as follows. One can apply the inverse isometric transformation on the coefficients c^*_j. Wang, Shangguan, Wu, and Guan (2013) provide a a very nice overview of this procedure. In their paper, they show that the process of inverting the isometric transformation requires two steps. For the transformed vector $w \in \mathbb{R}^{D-1}$, one first computes vector $v = (v_1, ..., v_D)$ defined:

$$v_k = \sum_{j=k}^{D} \frac{w_j}{\sqrt{j(j+1)}} - \sqrt{\frac{k-1}{k}} w_{k-1}, \qquad 1 \le k \le D$$

with $w_0 = w_D = 0$. Next, one computes $u \in \mathbb{S}^D$ as:

$$u_k = \frac{\exp(v_k)}{\sum_{i=1}^{D} \exp(v_i)}, \qquad \forall k$$

Morais, Thomas-Agnan, and Simioni (2017) show that for an observation t, the marginal effect

[22] In their paper, they show how this analysis can incorporate classical regressors as well, as opposed to using this technique solely with compositional explanatory variables. In their primary model, they use a dummy variable to control for auto scrapping incentives some manufacturers may have had in given years, however, the interpretations extend without issue to continuous variables.

of regressor Z_k on the j^{th} component of the dependent compositional variable is:

$$me(y_{jt}, Z_{tk}) = (\log c_{jk} - \sum_{m=1}^{D} y_{mt} \log c_{mk}) y_{jt}$$

where the j^{th} component of c is the j^{th} component of $ilr^{-1}(c_j^*)$, as defined above. As opposed to standard regression techniques where marginal effects in linear models or elasticities in log-linear models are directly interpreted from the parameters, here, marginal effects must be interpreted as functions of levels.

6.2 Analysis of Vehicle Fleet Composition

Using the techniques outlined in Section 6.1, it is possible to analyze the development of vehicle registrations in a compositional framework. There has been significant change in the make-up of vehicle registrations across states as well as at the national-level over the time frame of the data set. Figure 6.1 plots total vehicle registrations from 2000 to 2016 broken down by type. Over this time period, total registrations in the US have increased by 20% from 226 million to almost 270 million, despite the population increasing by only 14%. It is apparent that the share of trucks has increased drastically during this period. Interestingly, not only have passenger vehicles declined in share, but the raw number of automobiles has decreased. This suggests some level of substitution between vehicles defined as passenger vehicles to those defined as trucks on the part of households, and not only a structural change in economic activity. As was mentioned in Section 3, Trucks is a broad category and includes vehicles that are to an extent substitutes with traditional passenger vehicles. This is one reason why the role of Trucks in the model of Section 5 cannot be mapped perfectly to the empirical role of Trucks uncovered in the data. For years 2000 to 2014, the Highway Statistics also makes available data on specific truck types. In their online publication, they caution the reader to be careful making cross-state comparisons with these categories as there is significant variation among each state's methodology for categorization by these specific sub-classes. The Federal Highway Administration has its own approach to aggregation of the state-provided data and does so with an effort to have as much consistency as possible. Figure 6.2 replicates the results of Figure 6.1 up to 2014, although with greater segmentation of the Trucks category. From the graph, it is evident that the massive expansion of the truck classification in registration composition comes from the large increase in rates of sports utility vehicle (SUV) ownership, which, on average, are much less fuel efficient than traditional automobiles.

I now show how the regression techniques described in Section 6.1 can be used on the compositional vehicle registration data. I convert registration data into shares by dividing each state-year's observations of motorcycles, autos, trucks, and buses by the total quantity of registrations for that state-year. This yields a 816×4 matrix of compositional data dependent variables, $\mathbf{Y}$, where each row is an observation belonging to the simplex, $\mathbb{S}^4$. On each row, I perform the isometric logarithmic transformation and obtain a 816×3 matrix, $\mathbf{Y}^*$, where each row is a transformed observation in $\mathbb{R}^3$. In my primary specification, I regress each of

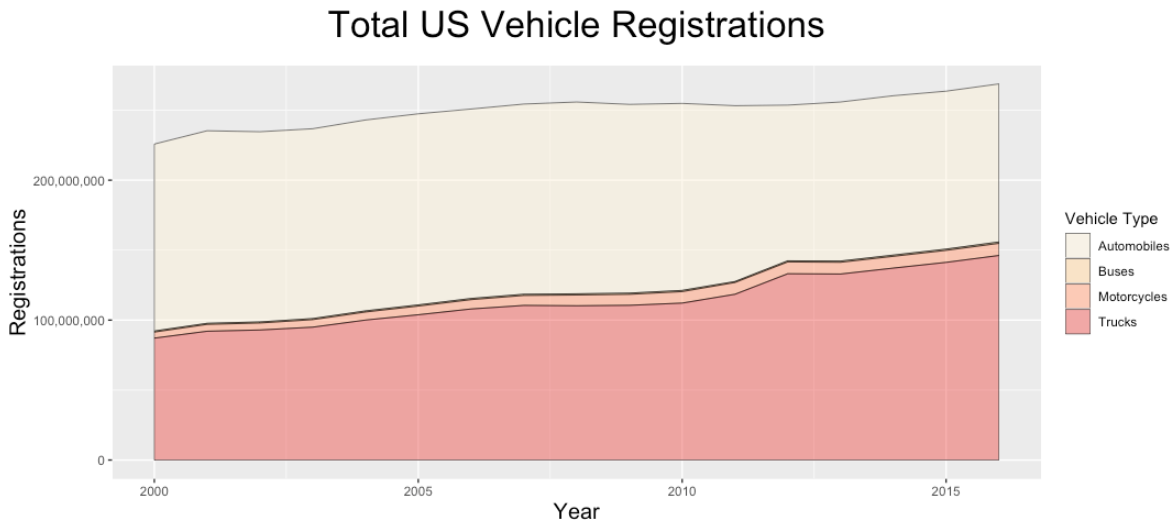

Figure 6.1: Vehicle Registration by type

these vectors of components on several covariates: gasoline tax rate, tax-exclusive retail price, GDP/capita, political structure integer variable (as defined in Section 3), public road length per square mile, and a time trend. The regression model is thus given by:

$$Y_{st}{}^{*} = a^{*} + B^{*} Z_{st} + c^{*} \mathrm{T}_{st} + \epsilon^{*}_{st}$$

where $Y_{st}{}^{*}$ is a three-dimensional vector of transformed data for state s and year t. a^{*} is a three-dimensional vector of constants. Z_{st} is a 5-dimensional vector of covariates as defined above. B^{*} is a 3×5 matrix of coefficients. Table 6.1 shows in the first column the results for the coefficients on gasoline tax; i.e. the elements of $c^{*}_{j} \in \mathbb{R}^{3}$. T-statistics and accompanying p-values indicate highly significant effects of the gasoline tax from this multivariate linear regression. The fourth column reports the transformation $ilr^{-1}(c^{*}_{j})$. These values are used to compute marginal effects with the Morais et al. (2017) equation.

Table 6.1: Result from isometric regression

Coefficient for Gasoline Tax				
	c^{*}_{j}	t-statistic	p-value	c_j
1	-0.879**	-4.22	< 0.001	**0.08368**
2	-0.846***	6.319	< 0.001	**0.1838**
3	-1.237***	-4.621	< 0.001	**0.1214**
4	–	–	–	**0.0581**

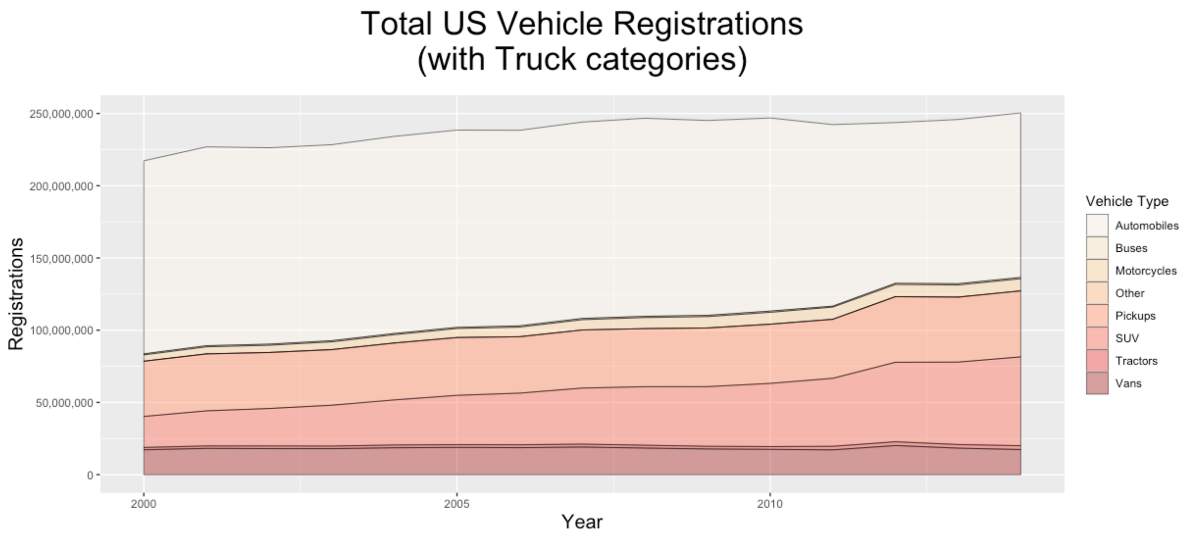

Figure 6.2: Vehicle Registration with truck categories

Because these marginal effects are a function of current shares, I calculate the effect an increase in the gasoline tax rate would have on the vehicle fleet composition for each state using the shares from 2016. As discussed in Section 1, Parry and Small (2005) find that U.S. tax levels are far below the optimal level. They suggest an optimal tax rate of $1.01 per gallon for the United States. Adjusting for inflation, this is equivalent to a per-gallon tax of $1.25 in 2016. For each state, I measured the marginal effect from an increase in the tax rate equal to the increase needed to achieve a tax rate of $1.25. The average percentage point change for each registration class in shares across all states are: increase of 4.2 points for motorcycles, increase of 6.2 points for automobiles, decrease of 10.1 points for trucks, and decrease of .3 points for buses. (Note that these values sum to 0.)

Table 6.2 reports the predicted change in vehicle registration percentages by type had each state raised its state tax to $1.066. (That is, $1.25 minus the federal tax rate of $0.184). As one would expect, the share of trucks decreases in every state, while the share of automobiles increases. This captures much of the substitution to passenger vehicles from trucks that can also serve the needs of passenger transportation needs, such as SUVs and vans. Motorcycle shares increase as well, representing the fact that this highly fuel efficient means of travel has a significant comparative advantage, under a higher gas tax. Interestingly, the share of buses decreases, although at rates very close to 0.

I take the Euclidean distance from 0 of each observation's predicted change in shares and find the five states whose composition is most affected by the tax increase are South Dakota, New Hampshire, Oklahoma, Indiana, and Alaska, while the five states whose compositions are least affected are Pennsylvania, Washington, North Carolina, Rhode Island, and California.

As a means of examining how these changes hold up with more segmented data, I perform the same regression technique described above for the data form 2000 to 2014, splitting the Truck category into its components: pick-ups, SUV's, tractors, and vans. The Tractor category refers to the cab unit of a big-rig. Most states have a specific registration for this part of the vehicle. Cargo containers are then attached. The regression results yield mostly significant coefficients, suggesting some of the development of the shares is explained by the gasoline tax, however, interpreting the coefficients in direct relation to individual components of the dependent variable is difficult owing to the differing dimensions between the two sample spaces.

Table 6.3 shows the predicted changes from the regression results for 2014 levels. Autos increase in share steadily in all states and SUV's decrease. It is reasonable to assume that this captures sum of the substitution effect on behalf of households, as they switch to more fuel efficient passenger cars. Pickups also decrease. Vans actually increase, possibly because SUV drivers they are easier to switch to for these consumers than passenger vehicles. The Other category includes a range of trucks that do not fit the other categories. This category is very small for every state-year and the predicted changes in share are almost 0.

The results from this analysis show that a steep increase in gasoline tax may spark a strong reaction in the composition of the transportation economy. Sudden tax changes in gasoline prices aimed at combating climate change will have effects that are of key interest to local policy-makers.

Table 6.2: Predicted changes from optimal tax level

	Δ Motorcycles	Δ Automobiles	Δ Trucks	Δ Buses
AL	0.0258	0.0774	-0.1022	-0.0010
AK	0.0590	0.0586	-0.1087	-0.0089
AZ	0.0363	0.0712	-0.1062	-0.0013
AR	0.0403	0.0641	-0.1010	-0.0033
CA	0.0281	0.0589	-0.0847	-0.0023
CO	0.0457	0.0623	-0.1060	-0.0020
CT	0.0351	0.0661	-0.0976	-0.0035
DC	0.0152	0.0758	-0.0795	-0.0116
DE	0.0335	0.0695	-0.0996	-0.0034
FL	0.0386	0.0603	-0.0961	-0.0028
GA	0.0285	0.0690	-0.0940	-0.0035
HI	0.0327	0.0754	-0.1060	-0.0021
ID	0.0385	0.0535	-0.0906	-0.0014
IL	0.0378	0.0700	-0.1050	-0.0028
IN	0.0465	0.0671	-0.1108	-0.0028
IA	0.0558	0.0485	-0.1025	-0.0018
KS	0.0428	0.0612	-0.1024	-0.0016
KY	0.0295	0.0675	-0.0951	-0.0019
LA	0.0368	0.0697	-0.1004	-0.0060
ME	0.0517	0.0527	-0.1015	-0.0029
MD	0.0311	0.0607	-0.0878	-0.0040
MA	0.0387	0.0650	-0.1018	-0.0019
MI	0.0390	0.0693	-0.1074	-0.0009
MN	0.0477	0.0562	-0.1011	-0.0027
MS	0.0177	0.0812	-0.0957	-0.0032
MO	0.0350	0.0740	-0.1050	-0.0041
MT	0.0445	0.0585	-0.1000	-0.0030
NE	0.0328	0.0634	-0.0912	-0.0050
NV	0.0367	0.0662	-0.1018	-0.0010
NH	0.0683	0.0505	-0.1170	-0.0018
NJ	0.0333	0.0781	-0.1076	-0.0038
NM	0.0433	0.0670	-0.1074	-0.0028
NY	0.0410	0.0643	-0.0995	-0.0058
NC	0.0244	0.0614	-0.0831	-0.0027
ND	0.0537	0.0528	-0.1040	-0.0024
OH	0.0425	0.0594	-0.0989	-0.0029
OK	0.0471	0.0668	-0.1132	-0.0007
OR	0.0394	0.0589	-0.0950	-0.0033
PA	0.0293	0.0435	-0.0699	-0.0028
RI	0.0360	0.0547	-0.0888	-0.0019
SC	0.0353	0.0750	-0.1067	-0.0035
SD	0.0977	0.0319	-0.1281	-0.0015
TN	0.0361	0.0697	-0.1019	-0.0040
TX	0.0220	0.0760	-0.0956	-0.0024
UT	0.0394	0.0593	-0.0968	-0.0019
VT	0.0544	0.0506	-0.1036	-0.0013
VA	0.0340	0.0764	-0.1065	-0.0040
WA	0.0298	0.0489	-0.0767	-0.0020
WV	0.0383	0.0540	-0.0910	-0.0013
WI	0.0621	0.0446	-0.1048	-0.0019
WY	0.0435	0.0526	-0.0925	-0.0036

Table 6.3: Predicted changes from optimal tax level

	Δ Auto	Δ Bus	Δ Motorcycles	Δ Tractors	Δ Pickups	Δ Vans	Δ SUV	Δ Other
AL	0.1024	-0.0010	0.0286	0.0047	-0.0699	0.0050	-0.0701	0.0001
AK	0.0847	-0.0045	0.0655	0.0056	-0.0782	0.0137	-0.0880	0.0011
AZ	0.0919	-0.0038	0.0444	0.0056	-0.0625	0.0067	-0.0824	0.0001
AR	0.0867	-0.0038	0.0440	0.0102	-0.0765	0.0068	-0.0676	0.0001
CA	0.0695	-0.0025	0.0284	0.0049	-0.0386	0.0035	-0.0652	0.0000
CO	0.0878	-0.0023	0.0506	0.0031	-0.0512	0.0069	-0.0950	0.0001
CT	0.0864	-0.0041	0.0374	0.0015	-0.0317	0.0043	-0.0942	0.0003
DC	0.0831	-0.0079	0.0157	0.0002	-0.0124	0.0013	-0.0829	0.0030
DE	0.0881	-0.0044	0.0392	0.0022	-0.0430	0.0064	-0.0890	0.0006
FL	0.0900	-0.0039	0.0471	0.0025	-0.0487	0.0050	-0.0921	0.0001
GA	0.1106	-0.0055	0.0368	0.0068	-0.0668	0.0079	-0.0899	0.0001
HI	0.0959	-0.0020	0.0409	0.0006	-0.0637	0.0089	-0.0814	0.0008
ID	0.0793	-0.0017	0.0476	0.0086	-0.0739	0.0070	-0.0677	0.0007
IL	0.0824	-0.0030	0.0432	0.0091	-0.0412	0.0059	-0.0965	0.0001
IN	0.0725	-0.0015	0.0484	0.0255	-0.0653	0.0055	-0.0852	0.0001
IA	0.0703	-0.0024	0.0660	0.0113	-0.0753	0.0061	-0.0764	0.0003
KS	0.0772	-0.0013	0.0479	0.0114	-0.0707	0.0063	-0.0712	0.0004
KY	0.0823	-0.0019	0.0308	0.0044	-0.0597	0.0064	-0.0623	0.0001
LA	0.0963	-0.0078	0.0398	0.0062	-0.0735	0.0063	-0.0673	0.0001
ME	0.0748	-0.0028	0.0502	0.0040	-0.0590	0.0054	-0.0731	0.0004
MD	0.0846	-0.0027	0.0380	0.0019	-0.0380	0.0048	-0.0888	0.0000
MA	0.0923	-0.0024	0.0308	0.0017	-0.0318	0.0056	-0.0961	0.0000
MI	0.0876	-0.0010	0.0415	0.0058	-0.0506	0.0085	-0.0917	0.0000
MN	0.0685	-0.0022	0.0524	0.0077	-0.0528	0.0054	-0.0791	0.0000
MS	0.1049	-0.0036	0.0187	0.0110	-0.0740	0.0067	-0.0643	0.0005
MO	0.0914	-0.0011	0.0391	0.0090	-0.0681	0.0076	-0.0780	0.0001
MT	0.0364	-0.0027	0.1334	0.0086	-0.0999	0.0015	-0.0778	0.0006
NE	0.0775	-0.0018	0.0331	0.0193	-0.0646	0.0064	-0.0704	0.0005
NV	0.0902	-0.0016	0.0399	0.0023	-0.0525	0.0053	-0.0836	0.0000
NH	0.0756	-0.0022	0.0741	0.0022	-0.0563	0.0038	-0.0976	0.0004
NJ	0.1040	-0.0038	0.0309	0.0043	-0.0290	0.0065	-0.1131	0.0001
NM	0.0926	-0.0035	0.0457	0.0059	-0.0773	0.0065	-0.0705	0.0006
NY	0.0818	-0.0022	0.0374	0.0022	-0.0292	0.0059	-0.0959	0.0001
NC	0.0763	-0.0030	0.0255	0.0040	-0.0465	0.0059	-0.0622	0.0001
ND	0.0649	-0.0030	0.0570	0.0298	-0.0843	0.0062	-0.0710	0.0003
OH	0.0730	-0.0034	0.0445	0.0058	-0.0468	0.0047	-0.0779	0.0001
OK	0.0895	-0.0006	0.0493	0.0078	-0.0747	0.0071	-0.0784	0.0000
OR	0.0833	-0.0019	0.0304	0.0044	-0.0554	0.0072	-0.0682	0.0002
PA	0.0743	-0.0045	0.0423	0.0045	-0.0393	0.0050	-0.0824	0.0001
RI	0.0712	-0.0021	0.0395	0.0014	-0.0327	0.0026	-0.0804	0.0005
SC	0.0996	-0.0044	0.0391	0.0039	-0.0625	0.0069	-0.0826	0.0001
SD	0.0479	-0.0027	0.1088	0.0156	-0.0915	0.0025	-0.0814	0.0007
TN	0.0927	-0.0045	0.0396	0.0057	-0.0649	0.0071	-0.0758	0.0001
TX	0.0991	-0.0045	0.0290	0.0089	-0.0632	0.0070	-0.0763	0.0000
UT	0.0827	-0.0021	0.0430	0.0074	-0.0586	0.0058	-0.0783	0.0000
VT	0.0829	-0.0032	0.0645	0.0039	-0.0620	0.0047	-0.0910	0.0002
VA	0.1048	-0.0047	0.0386	0.0026	-0.0545	0.0077	-0.0946	0.0001
WA	0.0697	-0.0026	0.0374	0.0040	-0.0480	0.0048	-0.0655	0.0001
WV	0.0744	-0.0017	0.0410	0.0031	-0.0577	0.0057	-0.0654	0.0005
WI	0.0615	-0.0022	0.0586	0.0069	-0.0514	0.0045	-0.0780	0.0002
WY	0.0743	-0.0014	0.0510	0.0098	-0.0774	0.0073	-0.0645	0.0009

7 Conclusion

Despite being under-utilized when compared to most developed countries, gasoline taxes in the United States have substantial ability to affect how economic activity is conducted. While studies of household behavior in relation to gasoline price effects and tax levels are abundant, this paper utilized aggregate data to tell a broader story of economic reactions to this important tax mechanism. Measuring aggregate vehicle utilization by vehicle registrations, I found an elasticity with respect to tax of -0.208 in the preferred specification and a tax-exclusive price elasticity of -0.116. The output per capita elasticity was measured at 1.214, and is significantly greater than 1, implying a given economy's level of vehicle utilization is highly elastic with respect to output. With more detailed data using measurements of capital and labor, one can investigate this result further to measure how specific inputs of output determine this elasticity, such as total factor productivity, or hours worked. I also found that aggregate transportation demand, as measured by total vehicle miles travelled is much more responsive to tax changes than equivalent changes to the tax-exclusive price. This result provides interesting insight into consumer behavior and fuels the argument that tax polices may wield more power than price-based studies suggest.

Next, I developed a theoretical static model of discrete vehicle choice on behalf of consumers who also make decisions over labor, leisure, and consumption. The purpose of this model was to treat VMT in a new light by exogenizing this economic component and incorporating a production process that uses household labor, and uses polluting transportation capital to produce goods. This qualification on VMT reflects the fact that households generally are not able to move their residence or find new employment opportunities. Government establishes taxation policies and a local entity carries out policy proposals. The interaction of the various economic agents demonstrates that transportation is an intricate aspect of economic behavior and that taxation can potentially have far-reaching consequences. While the model is simplified, a full development lies out of the scope of what this paper attempts to study. Nonetheless, this theoretical insight offers an interesting framework through which to view the econometric results of the concluding section.

Using segmented data describing vehicle registrations, I used statistical tools for analyzing compositional data, to model how composition of the vehicle fleet responds to changing taxation policies. I found evidence that gasoline taxation has a strong impact on the dynamics of this composition. Specifically, following a hike in the gasoline tax to the optimal level as outlined by Parry and Small (2005), one should anticipate a 4.2 percentage point increase in the share of motorcycles, a 6.2 point increase in the share of automobiles, a 10.1 point drop in trucks, and a 0.3 point drop in buses.

A thorough analysis of data within a compositional framework can offer important insights to studying economic behavior. While standard regression techniques on raw levels is certainly interesting, studying proportional outcomes explicitly can provide a new perspective, especially in the field of taxation. As economic studies increasingly incorporate environmental models, robust analyses of consumer behavior are crucial.

References

Agras, J. and Chapman, D. (1999). The Kyoto Protocol, Cafe Standards, And Gasoline Taxes. *Contemporary Economic Policy*, 17(3):296–308.

Aitchison, J. (1986). The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):139–160.

Allcott, H. and Wozny, N. (2014). Gasoline Prices, Fuel Economy, and the Energy Paradox. *The Review of Economics and Statistics*, 96(5):779–795.

Austin, D. and Dinan, T. (2005). Clearing the air: The costs and consequences of higher cafe standards and increased gasoline taxes. *Journal of Environmental Economics and Management*, 50(3):562–582.

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

Busse, M. R., Knittel, C. R., and Zettelmeyer, F. (2013). Are Consumers Myopic? Evidence from New and Used Car Purchases. *American Economic Review*, 103(1):220–256.

Dahl, C. A. (1986). Gasoline demand survey. *The Energy Journal*, 7(1):67–82.

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

Doyle Jr., J. J. and Samphantharak, K. (2008). $2.00 Gas! Studying the effects of a gas tax moratorium. *Journal of Public Economics*, 92(3-4):869–884.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3):279–300.

Goldberg, P. K. (1998). The effects of the corporate average fuel efficiency standards in the us. *The Journal of Industrial Economics*, 46(1):1–33.

Golob, T. F., Beckmann, M. J., and Zahavi, Y. (1981). A utility-theory travel demand model incorporating travel budgets. *Transportation Research Part B: Methodological*, 15(6):375 – 389.

Hausman, J. A. and Newey, W. K. (1995). Nonparametric estimation of exact consumers surplus and deadweight loss. *Econometrica*, 63(6):1445–1476.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, 39(153):41–57.

Klier, T. and Linn, J. (2010). The price of gasoline and new vehicle fuel economy: Evidence from monthly sales data. *American Economic Journal: Economic Policy*, 2(3):134–153.

L. McFadden, D. (1974). *Conditional Logit Analysis of Qualitative Choice Behavior*, volume 8, pages 105–142.

Li, S., Linn, J., and Muehlegger, E. (2014). Gasoline taxes and consumer behavior. *American Economic Journal: Economic Policy*, 6(4):302–342.

Li, S., Timmins, C., and von Haefen, R. H. (2009). How do gasoline prices affect fleet fuel economy? *American Economic Journal: Economic Policy*, 1(2):113–137.

Mayo, J. W. and Mathis, J. E. (1988). The effectiveness of mandatory fuel efficiency standards in reducing the demand for gasoline. *Applied Economics*, 20(2):211–219.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3(4):303–328.

Morais, J., Thomas-Agnan, C., and Simioni, M. (2017). Impact of advertising on brand's market-shares in the automobile market: a multi-channel attraction model with competition and carryover effects. working paper.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 114(7):1518–1523.

Parry, I. W. H. and Small, K. A. (2005). Does britain or the united states have the right gasoline tax? *The American Economic Review*, 95(4):1276–1289.

Pearson, K. (1897). Mathematical contributions to the theory of evolution. 2014; on a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, 60(359-367):489–498.

Wang, H., Shangguan, L., Wu, J., and Guan, R. (2013). Multiple linear regression modeling for compositional data. *Neurocomput.*, 122:490–500.

West, S. E. and Williams, R. C. (2004). Estimates from a consumer demand system: implications for the incidence of environmental taxes. *Journal of Environmental Economics and Management*, 47(3):535 – 558. Including Special Symposium Section from the National Bureau of Economic Research Conference on Advances in Empirical Environmental Policy Research.