Income Elasticity of Driving and Regressivity of Emission Control Taxation Policies: Evidence from Massachusetts New Car Market

Tingmingke Lu*

(PRELIMINARY AND INCOMPLETE: PLEASE DO NOT CITE OR DISTRIBUTE)

Abstract

What is the distributional impact of emission control taxation policies on new car markets? This paper examines the relative regressivity of taxes on fuel and taxes on the low fuel economy of cars, conditioning on car purchases. To estimate how tax revenues change with income, I rely on nameplate-model fixed effects indicated in a rich panel of new cars sold in Massachusetts and exploit the correlation between driving patterns and average demographic characteristics at the ZIP Code level to identify flexible substitution patterns for new car buyers. The heterogeneity allowed in my model demonstrates that the difference between the elasticities of tax revenues with respect to income across the two policies lies in the income elasticity of driving. Theory shows that if the income elasticity of driving is positive, the elasticity of fuel tax revenues with respect to household income must be greater than that of taxes on low fuel economy. Policy counterfactuals confirm the theoretical result. Although wealthier people are less likely to purchase more fuel-efficient cars when facing trade-offs between vehicle fuel efficiency and other vehicle attributes, the fact that more affluent households drive more makes fuel taxes more progressive relative to taxes on low fuel economy. Moreover, I show that the loss in total consumer surplus from applying fuel taxes is smaller than that from applying taxes on low fuel economy for achieving the same level of externality reduction because fuel taxes have heterogeneous influences on consumers in different mileage groups. Therefore, this paper provides a framework for understanding the distributional consequences of regulating the market of energy-using durable goods with various policy instruments in the presence of usage heterogeneity.

*Email: tingmingke.lu@slu.se. Web: https://sites.google.com/view/tingmingkelu
1. Introduction

Emissions from road transport make up the most significant share of US carbon emissions from fossil fuel consumption among all end-use economic sectors (EPA, 2018). In the meantime, tailpipe emissions generate externalities that lead to high health care costs (Knittel et al., 2016; Simeonova et al., 2019; Wu et al., 2020). With growing concerns about energy overuse, climate change, and the impact of local air pollution on public health, federal and state governments have been working on moving buyers toward more fuel-efficient vehicles and accelerating the adoption of low-emission vehicles to reduce fleet emissions.\(^1\)

Fuel economy taxation policies that tax or subsidize new vehicle purchases based on vehicle fuel economy performance have played an influential role in the US for curbing fuel consumption and reducing emissions from road transport. However, Borenstein and Davis (2016) show that tax expenditures promoting hybrid electric vehicles (HEVs) and battery electric vehicles (BEVs) between 2006 and 2012 have gone predominantly to higher-income Americans. Suppose the goal is to facilitate the transition toward a more sustainable and equitable road transport system—it makes sense for policymakers to consider the distributional implications of alternative policies as well as their political expediency.

Economists have been comparing fuel economy taxation policies with fuel taxes on efficiency grounds. But there haven’t been many works comparing the distributional effect of these two types of policy instruments.\(^2\) A handful of studies looking into the relative regres-

---

\(^1\)For instance, a goal that makes all new cars sold be electric by 2040 has been recommended to the Baker-Polito Administration to address future transport needs and challenges in Massachusetts. Refer to this WEB PAGE for more details about the Massachusetts Commission on the Future of Transportation (information retrieved on July 17, 2020). This US Department of Energy WEB PAGE provides more information on new and recently updated state laws, incentives, and regulations related to environmentally friendly vehicles.

\(^2\)Sallee (2011b) suggests that direct fuel taxation will deliver larger benefit per dollar of social cost if the primary aim of a policy is to reduce fuel consumption, improve local air quality, and mitigate climate change. Deryugina et al. (2019) point out that if efficient policies have undesirable distributional consequences and if corrective distributional policies are not feasible, distributional and efficiency trade-offs may advocate the use of policies that do not directly price externalities, such as fuel economy taxation policies.
sivity put aside the price effect of fuel economy taxation policies. Taxes and subsidies tied to the fuel economy rating of cars lead consumers to make trade-offs between vehicle fuel efficiency and other vehicle attributes. Because a change in fuel economy is usually bundled with changes in other car attributes, such trade-offs may alter the transportation service the car provides (Gillingham et al., 2016). For instance, a consumer who bought a subsidized vehicle might not be a winner because she would have purchased another model with a specific feature that she prefers if there isn’t a subsidy linked to that car’s fuel economy performance (Davis and Knittel, 2019). Furthermore, prior studies based on the expenditure data have shown that rich people purchase more new cars, and they drive more. Consequently, they benefit more from subsidies while paying more fuel taxes, so fuel taxes are relatively less regressive (Levinson, 2019). However, this might not be true conditional on new car purchases because cars of the same model years are similar in fuel economy ratings. Wealthier households may not necessarily buy more fuel-efficient new cars, although they may still drive more. In such a case, more affluent households pay more fuel taxes but may not benefit more from subsidies linked to cars’ fuel economy performance. Therefore, the relative regressivity of the two policies is not clear from comparing static tax incidence.

In this paper, I model fuel economy taxation policies as taxes on the low fuel economy of cars calculated based on the amount of fuel consumed for driving one mile of a car model (measured in gallon-per-mile).\(^3\) By estimating a structural model of new car demand paired with a joint distribution of household demographic characteristics and mileage types, I estimate how tax revenues change with income. Letting consumers vary in their marginal utility of money and vehicle use patterns, the heterogeneity allowed in my model demonstrates that the difference in elasticities of tax revenues across the two policies equals the income elasticity of demand for driving. Through model derivations and policy counterfactuals, I show

\(^3\)Framing fuel economy taxation policies in this way makes them more comparable to an energy tax. See Fischer (2009) and Sallee (2011b) for thorough examinations of the design of US fuel economy taxation policies.
that if the income elasticity of driving is positive, the elasticity of fuel tax revenues with respect to household income must be greater than that of taxes on the low fuel economy of cars. This observation reveals that the targeting mechanism of the fuel tax makes it more progressive relative to fuel economy taxation policies in emissions reductions.

Past literature has pointed out that an individual’s car choice and driving choice are correlated because car attributes that influence the car purchase decision might be associated with that individual’s car usage decision. For instance, a new car buyer living far from her work location is more likely to buy a fuel-efficient car because she expects to drive more. Several works have made attempts to capture the expected driving known by the consumer at the time of the new car purchase but not the econometrician. West (2004) applies a “typical quarterly miles” component for households in each car bundle in her nested logit model for car choices. Bento et al. (2009) apply cross-equation restrictions to estimate car choices and driving decisions in a unified behavior model. Gillingham (2012) employs a structure that includes the “known utilization type” and the “unknown preference for driving”, both of which are drawn from Normal distributions.

Similar to Grigolon et al. (2018), I employ the empirical distribution of observed mileage to characterize driving types. I rely on spatial heterogeneity and compute the ZIP Code average annual mileage of new cars as in Nurski and Verboven (2016). Therefore, this paper uses the nameplate-model fixed effects indicated in a rich panel data of new cars sold in Massachusetts and exploits the correlation between driving patterns and the average demographic characteristics at the ZIP Code level to identify flexible substitution patterns for new car buyers.

This paper relates to several strands in the literature. Davis and Knittel (2019) quantify the magnitude of implicit subsidies and taxes imposed by the US Corporate Average Fuel

---

4Different from their implementation, I don’t need the unconditional fractions of mileage types. In my model, new car buyers choose between gasoline-powered cars and HEV/BEV instead of buying a car or not, so fractions of observed mileage conditional on purchasing a new car suit my purpose.
Economy (CAFE) standards. They show that the fuel economy standards are mildly progressive when considering new vehicles only. High-income households bear more of the cost as a fraction of income than low-income households, which largely reflects that high-income households buy more new vehicles. However, their results come from matching vehicle registrations with demographic information at the census tract level, assuming taxes and subsidies do not change households’ car choice decisions.

Using the Consumer Expenditure Survey (CEX), West (2004) compares the distribu­tional consequences of gasoline taxes to those of a tax on vehicle size. She indicates that a tax on size is significantly more regressive than a gasoline tax. Applying updated CEX data, Levinson (2019) points out that high-income households consume more energy services and own cars with better fuel economy ratings. Therefore, high-income households pay more energy taxes in absolute terms while benefiting from the energy efficiency standards more, making energy efficiency standards more regressive. Although informative, this set of studies analyzes the static incidence of taxes and abstracts from consumer responses to policy changes.\(^5\) In contrast, the present paper adopts a structural approach that allows for conducting counterfactual simulations to analyze the substitution behavior of consumers in the new car market under alternative incentive schemes.

Durrmeyer (2020) analyzes the simulated market equilibrium based on a structural model to investigate the impacts of the 2008 French feebate policy. She describes the gains and losses of households in terms of a direct price effect and an indirect effect from changes in local air quality. Although rich demographic information is applied to characterize the distributional impact of a feebate policy, an assumption that the annual mileage driven is fixed across all locations makes the link between policy effects and vehicle use heterogeneity less straightforward. The present paper models vehicle usage heterogeneity and policy influences in a unified framework to provide an intuitive interpretation of the distributional impacts of

\(^5\)West (2004) allows households to make updated mileage decisions conditional on their original car choices after applying new gasoline taxes.
alternative tax policies. Such an approach is in line with the suggestion made in Holland et al. (2019).

Employing a structural model to jointly estimate vehicle choice and usage, Bento et al. (2009) and Jacobsen (2013) have studied the distributional impacts of gasoline taxes and fuel economy standards separately. Bento et al. (2009) show that under a flat gasoline tax revenue-recycling scheme, the average household in each of the bottom four income deciles experiences a welfare gain from a gasoline tax increase, while if revenues are recycled in proportion to income, only very poor households and very rich households stand to gain. Jacobsen (2013) suggests that the fuel economy standards have a greater impact on wealthier households who purchase a disproportionate share of new vehicles. But in the long run, increased prices and changes in fleet composition for used cars lead to larger proportional welfare losses for low-income households. However, because of model complexity, this set of studies is restricted to characterizing vehicle choices by aggregated product categories (e.g., car classes, car age groups, and manufacturers), in which case a consumer’s decision-making on which vehicle to purchase is not well captured. The present paper builds upon a detailed demand estimation to provide a more in-depth understanding of trade-offs made by new car buyers between vehicle fuel efficiency and other vehicle attributes.

By applying the modeling framework developed in Nurski and Verboven (2016) to capture consumer heterogeneity in household income and car usage, this paper adds to recent literature employing techniques estimating consumer demand in markets of differentiated products to investigate the effects of emission control taxation policies on car purchase and car usage decisions. For example, Bento et al. (2012) and Gillingham (2012) show that new car buyers may self-select into different car categories according to anticipated driving. Grigolon et al. (2018) demonstrate that properly accounting for vehicle use heterogeneity helps characterize the effectiveness of fuel taxes. The present paper builds upon this important empirical literature but provides more insights into redistributive motives of policy designs instead of focusing purely on the effectiveness of policies.
This paper is also related to the theoretical literature in behavioral public economics that examines the design of optimal policies for addressing consumption internalities and externalities. Allcott et al. (2019) develop the theoretical derivation of an optimal tax that deals with the overconsumption of sin goods while accounting for both corrective and regressive concerns. Allcott et al. (2014) and Heutel (2015) suggest that it’s possible to develop a policy combination that consists of energy taxes and taxes on low fuel economy to deal with consumer undervaluation of energy costs. This article is similar to these works in terms of considering the targeting mechanism of alternative tax policies in the presence of vehicle use heterogeneity among new car buyers. The present paper differs from this set of studies by investigating the regressivity of separate existing policy instruments and focusing more on policy designs that address externalities.

While providing a better understanding of the distributional impacts of alternative emission control taxation policies in the new car market, the present work has some limitations. First, this paper models the compositional effect of policies but not the usage effect. By choosing to apply an inelastic demand for driving with respect to fuel prices, I assume that an individual’s expected annual mileage does not change before and after implementing new fuel taxes even if these taxes are likely to change the variable cost of driving.\footnote{In relation to this first limitation, it’s also reasonable to think that acquiring a fuel-efficient car may lead to more driving because the variable cost of driving becomes lower. Following the literature, I assume such a direct rebound effect is not significant. Refer to Gillingham et al. (2013) for a further discussion on the weak rebound effect.} Alternative to this inelastic demand for driving employed in this paper, a greater degree of price-responsiveness among low-income households mitigates the regressivity of a fuel tax.\footnote{Refer to West (2004) for estimated elasticities of demand for driving with respect to operating costs across different income groups.} Therefore, relaxing this assumption does not alter the policy implications of the paper.\footnote{A future version of this paper will have an analysis assuming differential driving responsiveness levels.}

Secondly, because this paper focuses on modeling consumer responses in a local new car
market over a short period, manufacturers’ adjustments are not considered. Therefore, the analysis carried out here only captures the short-run effect of policy changes on the demand side. Consequently, this paper assumes new car buyers fully absorb the influence of policy changes in the short run. Although this is a strong assumption, it does not depart from conclusions obtained in previous studies. For instance, in the US context, Sallee (2011a) uses transaction-level microdata to show that consumers fully capture both federal and state tax incentives for the Toyota Prius. In a recent study, Muehlegger and Rapson (2018) show that the rate of pass-through to buyers is indistinguishable from 100 percent for California’s retire-and-replace subsidy program for electric vehicle purchases. Finally, this paper is silent on how additional revenue from the tax is used.

The rest of the paper proceeds as follows. Section 2 describes the data. Section 3 demonstrates empirical observations and descriptive evidence. Section 4 introduces a structural model of new car demand and presents results from a demand estimation. Section 5 conducts policy counterfactuals. Section 6 concludes.

2. Background and data

The primary data used in this paper is the Massachusetts Vehicle Census (MAVC). It is based on the Automated License and Registration System and a separate database containing records of vehicle inspections, both of which are administrative data sets maintained by the Massachusetts Registry of Motor Vehicles (MAPC, 2015).

Vehicle data. In Massachusetts, passenger vehicle registration renewal is valid for two years while vehicles are required to be inspected annually and within seven days of sale. When constructing the vehicle census, registration records are split where an inspection record—which delivers the mileage reading—begins or ends. So each record in the MAVC covers

9Car manufacturers may respond to policy changes by advancing technological innovations. Refer to Klier and Linn (2010) and Knittel (2011) for discussions on such mid-run effects.
a defined period when the specified vehicle had a unique combination of owner, garaging address, and average daily mileage (Reardon et al., 2016).\textsuperscript{10}

This paper applies highly disaggregated definitions of vehicle models to capture the variation in fuel efficiency and engine performance as much as possible. Each vehicle recorded in the MAVC is a nameplate–model level manufacturer/model/model year combination (e.g., Volkswagen Jetta 2011). I use vehicle identification numbers to define trim models in this analysis.\textsuperscript{11} The unit of observation in my sample is at the very detailed trim–model level (e.g., Volkswagen Jetta 2011, 2.5 liters, 170 hp, 3,045 lbs, 27 MPG, 94.1 ft\textsuperscript{3} passenger volume, and 15.5 ft\textsuperscript{3} cargo volume).

**Monthly new car market.** Metropolitan Area Planning Council Data Services Department suggests that a complete enumeration of registered vehicles in the vehicle census starts since the calendar year 2011.\textsuperscript{12} I also avoid the model year 2010 car models and previous model years because of the automobile industry crisis caused by the great recession from 2008 to 2010. After excluding earlier observations, new passenger cars registered in the second half of the calendar year 2011 have the most extended inspection history within the MAVC time frame for me to trace mileage records. I use new car registration records of the model year 2011 and the model year 2012 in Massachusetts between July 2011 and

\textsuperscript{10}Vehicle manufacturer, model, fuel type, fuel economy rating, curb weight, and the manufacturer suggested retail price (MSRP) of each vehicle are included in the MAVC data. Vehicle Identification Number (VIN) and ZIP Code of garaging address are also available from the MAVC researcher files.

\textsuperscript{11}Manufacturers use trim models to identify a vehicle’s level of equipment or special features. For models that use several trim choices, automakers usually offer three or four versions. For example, the gasoline-powered 2011 Volkswagen Jetta comes in three versions: S, SE, and SEL. The Jetta S is the base model, which includes the fewest features and has the lowest price of the three. The SE is in the middle of the range in both price and equipment, and the SEL is the most luxurious and feature-rich version. I use the VIN decoder provided by the National Highway Traffic Safety Administration to retrieve the trim level information of each vehicle. Then I use the trim level information to collect extra vehicle attributes from Cars.com and Ward’s Automotive Yearbook.

\textsuperscript{12}The number of vehicle registrations in the MAVC was 25\% and 32\% lower than the number of registered vehicles published by the Massachusetts Department of Revenue for 2009 and 2010.
December 2011 to construct the monthly new car market sample.\textsuperscript{13}

I employ MSRPs as new car purchase prices because transaction prices are not available.\textsuperscript{14} Copeland et al. (2011) have shown that new car dealers drop the price of a particular vintage until the introduction of the next vintage of the same model in early summer each year. Consequently, a consumer may time her new car purchase according to dealers’ marketing plans. The use of monthly new car markets is designated to address such concerns. I include month-of-the-year market fixed effects in the demand estimation to account for the timing of purchases.

**Fuel prices.** Based on the Michigan Survey of Consumers (MSC), Anderson et al. (2011) conclude that households typically form expectations about the inflation-adjusted price of gasoline using a simple no-change model.\textsuperscript{15} Although the inflation expectation data in the MSC is limited to selected horizons, Baumeister and Kilian (2016) show this no-change model is valid when advanced methods of inflation forecasting are employed. I adapt this simple rule of thumb as the consumer gas price expectation.

Furthermore, I choose to model the fuel price expectation as not varying across new car buyers in different monthly markets. In particular, I let the fuel price expectation employed in my analysis be the average of the statewide average gasoline price in Massachusetts over the six months from July 2011 to December 2011. To avoid invoking the assumption of symmetric consumer responses to changes in fuel prices and vehicle fuel efficiency levels, I

\textsuperscript{13}Based on disaggregated new vehicle transaction data, Copeland et al. (2011) show for about half the calendar year, automakers simultaneously sell two vintages of the same model. In MAVC, less than 1\% of newly registered passenger cars during the second half of the calendar year 2011 are for the model year 2010. Refer to Appendix B.2 for the detailed data cleaning process.

\textsuperscript{14}The use of transaction prices gives limited improvements even if it captures individual-specific discounts. Nurski and Verboven (2016) suggest that even with household data, researchers observe the transaction price of the chosen alternative but not the ones of the non-chosen alternatives.

\textsuperscript{15}Specifically, households surveyed expect the nominal gasoline price to grow at the same rate as inflation, which is equivalent to expecting the real price of gasoline in the future to be the same as the current price of gasoline.
leave aside variations in fuel prices.\textsuperscript{16}

**ZIP Code level household income and annual new car mileage.** It is reasonable to expect that the distribution of consumer demographics affects new car sales. Therefore, I relate the information on household demographics to car purchases. I obtain median household income by ZIP Code Tabulation Area (ZCTA) from the 2011 American Community Survey (ACS). Because the ZIP Code of the garaging address is available for each vehicle in the MAVC, I apply the ZIP Code to ZCTA Crosswalk provided by the Uniform Data System (UDS) Mapper to map the ZCTA median household income to new car registrations in the sample.

I construct the ownership and inspection history for observations in the monthly new car market sample using the MAVC vehicle registration and inspection records. Each valid vehicle inspection record in the vehicle census reports the number of days between two inspections and the average daily miles traveled during that period. I weight the daily miles using the length of the period between two inspections and apply this weighted average daily mileage to calculate the expected annual vehicle miles traveled. About 70\% of the vehicles included in the new car market demand sample have three consecutive inspection records from 2011 through 2014.\textsuperscript{17} These observations are used for extracting average annual new car mileage at the ZIP Code level.

\textsuperscript{16}Linn (2016) and Gillingham (2020) point out this assumption may not hold in practice for a variety of reasons. For instance, Gillingham et al. (2015) find substantial heterogeneity in consumer response to gasoline prices by vehicle fuel economy quantiles, which implies that consumers may value a decrease in fuel price and an increase in vehicle fuel efficiency differently because the former can be seen as a relatively short-term gain in comparison to the latter.

\textsuperscript{17}Because a yearly inspection is required by law for all vehicles driven on Massachusetts roads, and the end-of-year sample cutoff causes a part of missing inspection records, I treat the rest of the missing records as random.
3. Empirical observations and descriptive evidence

This section describes the sample constructed from the MAVC. The present paper mainly focuses on the market for new passenger cars because most HEV/BEV models were concentrated in the passenger car category during the sample period.\textsuperscript{18} Therefore, this analysis excludes light-duty trucks, including pickup trucks, minivans, and sport utility vehicles.\textsuperscript{19}

**Summary statistics.** Panel A of table 1 describes the number of unique models observed in each monthly new car market. In total, there are 75 distinct nameplate models in the sample. Most of them repeatedly appear across six monthly new car markets. In the meantime, within each nameplate–model group, the trim–model combination defined using VINs and model years varies over time. Recognizing the panel feature of my sample, I employ nameplate–model fixed effects and car attributes that vary across trim models to provide identification in a demand estimation.

Panel B of table 1 reports on product differentiation in the new passenger car market. Although a few vehicle attributes have been collected, I only include four of them because many of the other variables are correlated. The literature studying automobile demand also suggests that vehicle performance and size are top attributes to consider when US consumers buying new cars.

The bottom panel of table 1 summarizes the annual household income and mileage of new car buyers at the ZIP Code level. I employ the method of simulated moments to estimate the demand system. To do so, I obtain a synthetic new car buyer population in each monthly new car market using the corresponding empirical distribution of demographics constructed from observed purchases. Applying the data that provides about 1,500 purchases made on 260 unique trim models each month in Massachusetts, I match simulated market shares to

\textsuperscript{18}Refer to table B.4 for more information about HEV/BEV models in the sample. Diesel-powered cars are not included in this analysis because their total market share was less than 1% during the sample period.

\textsuperscript{19}Refer to table B.5 for more information about model types in the MAVC sample.
observed market shares. Although local level market shares are available (e.g., ZIP Code area and municipality), I match market shares in the statewide market because there are too few trim–model level sales observed in separate local markets.

Table 1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Median</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Number of unique models in each monthly new car market</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unique Nameplate Model</td>
<td>70</td>
<td>3</td>
<td>72</td>
<td>6</td>
</tr>
<tr>
<td>Unique Trim Model</td>
<td>258</td>
<td>16</td>
<td>263</td>
<td>6</td>
</tr>
<tr>
<td><strong>Panel B: Trim-model level gas-powered passenger car</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSRP ($)</td>
<td>23,151</td>
<td>6,938</td>
<td>23,000</td>
<td>503</td>
</tr>
<tr>
<td>Performance (Horsepower/Weight)</td>
<td>0.540</td>
<td>0.107</td>
<td>0.512</td>
<td>503</td>
</tr>
<tr>
<td>Passenger Volume (100 ft³)</td>
<td>0.984</td>
<td>0.083</td>
<td>0.970</td>
<td>503</td>
</tr>
<tr>
<td>Fuel Economy Rating (MPG)</td>
<td>28</td>
<td>6</td>
<td>27</td>
<td>503</td>
</tr>
<tr>
<td><strong>Panel C: Household information by ZIP Code</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Income ($)</td>
<td>75,779</td>
<td>28,496</td>
<td>73,358</td>
<td>501</td>
</tr>
<tr>
<td>Annual Mileage (miles)</td>
<td>13,805</td>
<td>4,458</td>
<td>13,138</td>
<td>501</td>
</tr>
</tbody>
</table>

Table 1: Monthly new car markets during the second half of calendar year 2011.

Descriptive evidence. Just as summary statistics presented in table 1 provide information for identifying the heterogeneity parameters in a demand estimation, so table 2 demonstrates the correlation/a lack of correlation that characterizes the distributional impacts of tax policies in the Massachusetts new car market.

Levinson (2019) suggests that if the service provided by energy—miles of travel—is “normal” in an economic sense, then wealthier people want more of it. Applying all registered new cars with complete inspection records over three consecutive years, the first two columns in table 2 suggest that the income elasticity of driving estimated using vehicle-level annual mileage is about 0.15, meaning that a 10% increase in household income is associated with a 1.5% increase in new car usage.

However, the next two columns in table 2 suggest that more affluent households are not necessarily purchasing more fuel efficiency units in the new car market. In fact, conditional on other car attributes, wealthier families buy less fuel-efficient cars. This contradicts obser-
vations from the CEX in Levinson (2019). A possible reason could be that when looking at car stocks including both new and used cars, richer people purchase more fuel efficiency units because they own more new vehicles equipped with advanced fuel-saving technologies. The positive relationship between household income and car fuel efficiency disappears in the new car market because car models of the same model year are similar in fuel economy ratings.

Regression analyses presented in table 2 reflect the correlation between new car buyers’ household income and new car driving patterns. I leverage this empirical fact to set up a random utility model incorporated with consumer heterogeneity. To better understand the demand-side trade-offs made by consumers between fuel efficiency and other vehicle attributes in the new car market and their implications for the regressivity of alternative emission control tax policies, I turn to a structural model in the next section.

4. Demand for new passenger cars

In this section, I follow McFadden et al. (1973) to formulate a random utility model in which consumers choose between a set of alternative cars. Applying a discrete choice framework, I define the probability of a consumer buying a specific new car model as a function of the car’s attributes (including price), demographic characteristics, and the present discounted value of the car’s lifetime fuel cost based on this consumer’s expected annual mileage. In such a setup, cars with higher fuel economy ratings become relatively cheaper when taxes on low fuel economy are applied. Fuel taxes influence new car buyers’ vehicle purchase decisions by augmenting the difference in the fuel cost between vehicle models with different fuel economy ratings. Therefore, both taxes can incentivize new car buyers to choose fuel-efficient vehicles. To account for consumer heterogeneity, I let the preference for both cost components vary with household income and new car usage at the ZIP Code level.

4.1. A model of new car demand

I start from a random coefficients logit model of individual choice to obtain an aggregate demand system for differentiated new passenger cars as in Berry et al. (1995) and Nevo
Table 2: Household Income, New Car Annual Mileage, and New Car Fuel Efficiency

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>ln(Annual Mileage)</th>
<th>MPG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>ln(Household Income)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Performance (Horsepower/Weight)</td>
<td></td>
<td>-10.656</td>
</tr>
<tr>
<td>Passenger Volume</td>
<td>5.258</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>7.415</td>
<td>7.385</td>
</tr>
<tr>
<td></td>
<td>(0.577)</td>
<td>(0.592)</td>
</tr>
<tr>
<td>Trim-model FE</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Nameplate-model FE</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Model Year FE</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Municipality FE</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Monthly Market FE</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Number of Obs.</td>
<td>6,377</td>
<td>6,240</td>
</tr>
</tbody>
</table>

Table 2: Robust standard errors presented in parentheses are obtained by clustering samples at the municipality level. Results in the first two columns are from regressions using cars with three consecutive inspection records over three years. Such a mileage sample reports more accurate annual mileage estimates. Overall, the average number of days between two inspections is 379, and the standard deviation is 32. In a sample consists of cars with two inspections, these numbers are 408 and 110. For a mileage sample with cars having four inspection records, these numbers are 379 and 142.

(2001). Following Petrin (2002) and Nurski and Verboven (2016), I incorporate ZIP Code-level household information to facilitate the identification of the role played by consumer heterogeneity in the decision-making process.

When purchasing a new car, the consumer considers her utility from the car attributes and her disutility from the cost. Among all new car buyers observed in monthly market \( t \), a consumer living in ZIP Code area \( i \) choosing to buy a car of trim model \( j \) has a conditional indirect utility modeled as

\[
u_{ijt} = x_j' \alpha + \beta_i (p_j + \lambda h_{ij}) + \zeta_t + \zeta_{jt} + \epsilon_{ijt},
\]  

(1)
in which $x_j$ is a vector of observed car attributes that varies at the trim–model level, $p_j$ is the car purchase price, $h_{ij}$ is the expected annual fuel cost, $\zeta_t$ is the market specific common demand shock on all products, and $\zeta_{jt}$ is a product-market specific demand shock. I assume that $\epsilon_{ijt}$ is an individual specific taste parameter for car model $j$ in market $t$, modeled as a zero mean i.i.d. random variable with a Type I Extreme Value Distribution. Equation (1) models a consumer’s utility if she chooses to buy a gasoline-powered car. Instead of purchasing such a model $j$, this consumer may also consider to buy an HEV/BEV (i.e., the outside good). In that case, I specify $u_{i0t} = \epsilon_{i0t}$ to normalize the mean and individual specific valuations to zero because they are not identified from the constant.

In equation (1), $\alpha$ is the marginal utility from car attributes, and $\beta$ represents the marginal disutility from bearing the cost of purchasing the car and filling up the fuel tank for driving. I let the importance of both a car’s purchase price and the fuel cost be inversely related to household income. The assumption is that high-income households place less importance on such an expense. I therefore define

$$\beta_i = \frac{\beta}{y_i},$$

(2)

in which $y_i$ is the ZIP Code median household income.\(^ {20}\) I follow Grigolon et al. (2018) to specify the expected annual fuel cost as

$$h_{ij} = q e_j m_i,$$

(3)

in which $m_i$ is the expected annual new car mileage for a consumer living in ZIP Code area $i$, $e_j$ is the inverse of the fuel economy rating that indicates the vehicle fuel efficiency level of car model $j$ (i.e., $e_j = 1/mpg_j$, measured in gallon-per-mile), and $q$ is the fuel price

\(^ {20}\)Such a specification allows heterogeneity in the marginal utility of money to follow the empirical distribution of income. Refer to Grigolon and Verboven (2014) and Goldberg and Verboven (2001) for similar applications.
expectation that does not vary across new car buyers in different markets. The product of the fuel price and the vehicle fuel efficiency level indicates the dollar-per-mile value of a car model $j$ given $q$. The inclusion of term $\lambda$ in equation (1) reflects the fact that a consumer taking the present discounted value of future fuel costs into consideration when purchasing the car.\footnote{Parameter $\lambda$ can be modeled as a compound of a capitalization factor and a valuation factor as in Grigolon et al. (2018). The value of $\lambda$ is potentially associated with several dimensions of consumer heterogeneity and consumer valuation of expected future fuel costs, and is beyond the scope of this paper. Refer to Allcott and Wozny (2014) and Grigolon et al. (2018) for a detailed discussion.}

To incorporate emission control taxes into the model, I derive post-tax prices in terms of the original prices and the new tax components. I let $\tau^f$ denote the new fuel tax, which is an excise tax that changes fuel prices. A tax on low fuel economy is expressed as $e_j \tau^g$, which penalizes the purchase of fuel-inefficient cars by applying an extra cost unit tied to a car’s fuel efficiency level. Given these notations, the post-tax fuel price is $\tilde{q} = q + \tau^f$ per gallon and the post-tax vehicle purchase price becomes $\tilde{p}_j = p_j + e_j \tau^g$. Next, I combine equations (1) through (3) to get the individual choice probability

$$\tilde{s}_{ijt} = \frac{\exp(\tilde{\phi}_{ijt})}{1 + \sum_{j=1}^{J} \exp(\tilde{\phi}_{ijt})}, \quad (4)$$

in which

$$\tilde{\phi}_{ijt} = x_j' \alpha + \frac{\beta}{y_i} (\tilde{p}_j + \lambda \tilde{q} e_j m_i) + \zeta_t + \zeta_{jt}. \quad (5)$$

A model specified in this way allows me to examine the distributional impacts of taxes and interpret the income elasticity of tax revenues. In particular, given the individual choice probability described in equation (4), the probability weighted extra annual operating cost after implementing a new fuel tax is

$$T_{it}^f = m_i \tau^f \sum_{j=1}^{J} \tilde{s}_{ijt} e_j. \quad (6)$$
and the probability weighted extra price component of a car with the the tax on low fuel economy in effect becomes

$$T_{it}^g = \tau^g \sum_{j=1}^{J} \tilde{s}_{ijt} e_j. \quad (7)$$

Appendix A.1 shows

$$\frac{\partial T_{it}^f}{\partial y_i} \frac{y_i}{T_{it}^f} = y_i \sum_{j=1}^{J} \frac{\partial \tilde{s}_{ijt}}{\partial y_i} e_j + y_i \frac{\partial m_i}{\partial y_i} \quad (8)$$

and

$$\frac{\partial T_{it}^g}{\partial y_i} \frac{y_i}{T_{it}^g} = y_i \sum_{j=1}^{J} \frac{\partial \tilde{s}_{ijt}}{\partial y_i} e_j. \quad (9)$$

Comparing equations (8) and (9) gives

$$\eta_y^f = \eta_y^g + \eta_y^m, \quad (10)$$

in which $\eta_y^f$ is the income elasticity of the fuel tax revenue, $\eta_y^g$ is the income elasticity of the tax revenue for a tax on low fuel economy, and $\eta_y^m$ is the income elasticity of driving. Equation (10) indicates that the elasticity of the fuel tax revenue with respect to household income is larger than the income elasticity of the tax revenue of a tax on low fuel economy, with the difference between the two equalling the income elasticity of annual mileage.

Section 2 has shown that a 10% increase in household income is associated with a 1.5% increase in new car usage. As long as there is a correlation between household income and annual new car mileage, the distributional impacts differ for these two taxes. An analysis not accounting for consumer heterogeneity in household income and vehicle usage fails to demonstrate differentiated impacts of emission control taxation policies in the new car market.

4.2. A standard logit specification

The parameter of interest in equation (1) is $\beta_i$ because it governs the substitution behavior stimulated by changes in the cost components. To describe the identification of this parameter, I start with a logit model that does not account for any consumer heterogeneity.
**Specification.** I use the sample average income to scale prices and replace $m_i$ with the sample average annual mileage. The model specification becomes

$$u_{ijt} = x'_j \alpha + \frac{\beta}{y} (p_j + \lambda q e_j m) + \zeta_t + \zeta_{jt} + \epsilon_{ijt}$$

$$= x'_j \alpha + \beta \frac{p_j}{y} + \beta \lambda \frac{q}{y} e_j m + \zeta_t + \zeta_{jt} + \epsilon_{ijt}, \quad (11)$$

in which $\alpha$ and $\beta$ are separately identified while the product term $\beta \lambda$ is estimated as a single coefficient.

In equation (11), the only source of heterogeneity within each market is the individual-product specific shock $\epsilon_{ijt}$. The distributional assumption on $\epsilon_{ijt}$ leads to the following market share of model $j$ in market $t$

$$S_{jt} = \frac{\exp(x'_j \alpha + \beta \frac{p_j}{y} + \beta \lambda \frac{q}{y} e_j m + \zeta_t + \zeta_{jt})}{1 + \sum_{j=1}^J \exp(x'_j \alpha + \beta \frac{p_j}{y} + \beta \lambda \frac{q}{y} e_j m + \zeta_t + \zeta_{jt})}, \quad (12)$$

in which buying an HEV/BEV is characterized as the outside option with $j = 0$, $u_{i0t} = 0$, and $S_{0t} = 1 - \sum_{j=1}^J S_{jt}$. Following the transformation shown by Berry (1994), the difference between the logit market share for product $j$ and that of the outside option is

$$\log(S_{jt}) - \log(S_{0t}) = x'_j \alpha + \beta \frac{p_j}{y} + \beta \lambda \frac{q}{y} e_j m + \zeta_t + \zeta_{jt}. \quad (13)$$

**Identification.** With a linear transformation of equation (12), equation (13) can be estimated using OLS. Given the error term $\zeta_{jt}$, consistent estimation of $\beta$ for OLS is in doubt because prices are correlated with unobserved product attributes.\(^{22}\) To clarify the identification of $\beta$, I apply an instrumental variable (IV) approach estimated using two-stage least squares, and compare it with a fixed effect (FE) approach.

\(^{22}\)I assume that vehicle usage of new cars, the market segment considered here and which is likely to be a small fraction of total demand for driving, does not influence fuel prices.
First, I employ a set of price shifters $Z_{jt}$ proposed by Berry, Levinsohn, and Pakes (1995) for the IV approach. In equation (13), $\zeta_{jt}$ is assumed to be mean independent of $Z_{jt}$. Such demand-side IVs are developed from measures of isolation in the product space. They are commonly used when supply side IVs are not available. However, this approach assumes that unobserved and observed car attributes are not correlated, which is a strong assumption.

An alternative FE approach relies on nameplate–model fixed effects $\zeta_{j}^l$ to facilitate the identification of $\beta$. The source of identifying variation comes from the fact that the sample carries about 70 unique nameplate models over six monthly new car markets, and each nameplate model is associated with about three different trim models on average in each market as shown in Panel A of table 1. By exploiting the panel nature of the sample, this FE approach employs nameplate–model dummies to absorb unobserved product attributes. Then the specification becomes

$$log(S_{jt}) - log(S_{0t}) = x_j' \alpha + \beta p_j/y + \beta \lambda q e_j m/y + \zeta_{j}^l + \zeta_t + \zeta_{jt}. \quad (14)$$

4.3. A random coefficients logit specification

To incorporate consumer heterogeneity in household income and annual mileage into the estimation of new car demand, next I turn to a random coefficients logit specification of the model described in equation (1). Based on the empirical observations and descriptive evidence presented in section 3, such a modeling approach helps reveal the relative regressivity of alternative emission control taxes.

---

23In particular, for each car model, the squares of its own attributes, the sum of each attribute of car models produced by the same manufacturer, the sum of each attribute of car models made by competing manufacturers, and the number of unique models marketed by the same manufacturer are employed as instrumental variables. They are the standard instruments used in random coefficients logit demand applications and have been proved effective in the study of many industries such as automobiles, computers, and pharmaceutical drugs.
I restructure the conditional indirect utility function of equation (1) as

\[ u_{ijt} = \delta_{jt} + \mu_{ij} + \epsilon_{ijt}, \quad \text{with} \]

\[ \delta_{jt} = x_j' \alpha + \xi_j^t + \zeta_t + \zeta_{jt}, \quad \text{and} \]

\[ \mu_{ij} = \beta_i p_j + \beta_i \lambda q e_j m_i, \]

where \( \delta_{jt} \) is a trim-model specific mean utility component that does not vary with demographics, while \( \mu_{ij} \) contains the interactions between car attributes and household characteristics. Similar to Li (2012), I keep all cost components in the household-specific utility term \( \mu_{ij} \) and rely on nameplate-model fixed effects described in the previous subsection to identify coefficients in the mean utility term \( \delta_{jt} \). Given equation (15), the market share of car model \( j \) becomes

\[ S_{jt}(\delta_{jt}; \beta_i, \beta_i \lambda) = \int_{(y_i, m_i)} s_{ijt} dF(\beta_i, \beta_i \lambda), \]

in which

\[ s_{ijt} = \frac{\exp(\delta_{jt} + \mu_{ij})}{1 + \sum_{j=1}^J \exp(\delta_{jt} + \mu_{ij})}. \]

I estimate the model using the method of simulated moments and an iterative process. The moment condition is based on the exogeneity assumption that \( \zeta_{jt} \) is mean independent of \( x_j \) in equation (16). Given a synthetic new car buyer population, in each iteration I begin with a pair of \((\beta_i, \beta_i \lambda)\) to find a vector of \( \delta_{jt} \) that matches simulated market shares with observed market shares using a contraction mapping procedure proposed by Berry (1994) and Berry et al. (1995). Next, I construct the moment condition using \( \zeta_{jt} = \delta_{jt} - (x_j' \alpha + \xi_j^t + \zeta_t) \) and search for a full set of parameters to minimize \( E(\zeta_{jt}^t \zeta_{jt}) \).\(^{24}\)

\(^{24}\)I code analytical gradients to implement the nonlinear search. Using the quasi-Newton method and MultiStart function in MATLAB, I set the tolerance at 10e-14 for fixed-point iterations and test twenty starting values.
4.4. Estimation results

Table 3 presents estimation results from four demand model specifications. In column (1), when the correlation between $p_j$ and $\zeta_{jt}$ in equation (13) is not properly addressed, a linearized simple logit model estimated using OLS delivers a small and imprecise price coefficient. When combining a set of price shifters with a standard logit model, the estimation results presented in column (2) indicate that high cost reduces utility and new car buyers favor cars with large interior space.\textsuperscript{25} In column (3) of table 3, results produced from using an alternative identification strategy show comparable patterns of coefficient estimates. As described in equation (16), the random coefficients logit specification of the demand model exploits the panel nature of the sample and applies nameplate–model fixed effects to capture time-invariant unobserved attributes for a car. In column (4) of table 3, a random coefficients logit specification generates a set of consistent estimates for cost coefficients when consumer heterogeneity in household income and vehicle usage are added.\textsuperscript{26}

5. Policy counterfactuals

After obtaining estimates of coefficients that govern consumers’ substitution behavior in the demand system, I apply a fuel tax and a tax on low fuel economy to adjust the cost for new car buyers to purchase and drive a car. The comparison between outcomes with and without new taxes helps identify the causal effect of policy changes on the composition of new car sales. Unlike static analyses in prior studies, the estimation of a structural demand

\textsuperscript{25}To test the validity of this IV estimation, I regress prices on car attributes and corresponding price shifters. The F-statistic for the null hypothesis that coefficients on excluded instruments are zero in this OLS regression is 29, which allows me to reject the hypothesis at the 1% significance level. Armstrong (2016) suggests that the rejection of such a hypothesis implies using price shifters as instruments in this model is able to produce consistent estimates.

\textsuperscript{26}While the model specification requires me to keep all cost components in the household-specific utility term in the random coefficients logit case, columns (2) and (3) have shown that, in the standard logit case, the identifying power from applying nameplate–model fixed effects is as good as that from the price shifter IVs I can construct from the data.
<table>
<thead>
<tr>
<th></th>
<th>(1) Logit</th>
<th>(2) IV Logit</th>
<th>(3) FE Logit</th>
<th>(4) FE RC Logit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>S.E.</td>
<td>Est.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Fixed Coefficients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-5.67</td>
<td>0.36</td>
<td>-5.75</td>
<td>0.42</td>
</tr>
<tr>
<td>Car Performance</td>
<td>0.10</td>
<td>0.24</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>Passenger Volume</td>
<td>4.05</td>
<td>0.39</td>
<td>3.36</td>
<td>0.69</td>
</tr>
<tr>
<td>Annual Gas Cost/Avg. Inc.</td>
<td>-105.11</td>
<td>12.38</td>
<td>-72.76</td>
<td>21.52</td>
</tr>
<tr>
<td>Car Price/Avg. Inc.</td>
<td>-0.32</td>
<td>0.32</td>
<td>-1.76</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Coefficients, Interaction Terms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual Gas Cost/Inc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car Price/Inc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trim-model IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nameplate-model FE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly Market FE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Based on the specification in equation (13), Logit model in column (1) does not account for price endogeneity. IV Logit model in column (2) is estimated applying price shifters. FE Logit model in column (3) is estimated using the specification in equation (14). FE Random Coefficients Logit model in column (4) incorporates the interaction between cost components and observed demographics as presented in equation (17).

model conducted in this paper allows me to study the distributive nature of the amount of tax paid under different incentive schemes. In the meantime, the policy counterfactuals highlight distinguishable targeting mechanisms of alternative taxes by demonstrating changes in consumer surplus.

I compare tax revenues and changes in consumer surplus from implementing different taxes when both taxes are designed to increase the market share of HEV/BEV by the same amount. Specifically, I first compute the overall market share of HEV/BEV using the synthetic population simulated from data and their choice probabilities obtained from the demand estimation. Based on corresponding coefficient estimates reported in column (4) of table 3, I then solve for a new fuel tax that leads to a one-percent increase in the market share of HEV/BEV. Next, I solve for a new the tax on low fuel economy that leads to the
same change in the market share of HEV/BEV. Finally, I calculate and compare shares of taxes paid over the population sorted by income and changes in total consumer surplus when different taxes are in place. For policy counterfactuals presented in this section, a new fuel tax of $0.26/gallon and an alternative the tax on low fuel economy at purchase is implemented to increase the market share of HEV/BEV by one percent to 9.67%.  

In figure 1 (a), I use the relative concentration curve of the fuel tax with respect to the tax on low fuel economy to demonstrate the difference in distributional impacts of these two taxes. This visualization tool is generalized from the Lorenz Curve by Kakwani (1977). The horizontal axis of this figure ranks the synthetic new car buyer population by household income, while the vertical axis marks the cumulative share of tax paid. The 45-degree line represents a uniform tax on each household regardless of income. In such a figure, a lower curve indicates a less regressive tax scheme. Because poor households pay less tax than

\footnote{This number is comparable to that used in Levinson (2019), which equals $0.29/gallon. Refer to table C.6 and table C.7 for full schedules of taxes associated with different levels of HEV/BEV market share change.}

\footnote{Recent applications of this concentration curve on distributional effects of emission control taxation policies include Levinson (2019), Borenstein (2017), and Sterner (2012).}
wealthier households, lines representing both a fuel tax and a tax on low fuel economy are bowed downward. However, the cumulative share of the the tax on low fuel economy paid by households sorted by income rises faster than that of the fuel tax when both taxes are designed to expand HEV/BEV adoptions by the same amount. Therefore, up until the 80th percentile of the sorted population, the the tax on low fuel economy curve is closer to the 45-degree line than the fuel tax curve, meaning that a tax on low fuel economy is more similar to a household-level head tax in terms of regressivity. Kakwani (1977) also notices that the lower the position, the greater the elasticity with respect to the variable plotted on the horizontal axis. This interpretation is consistent with the result demonstrated in equation (10).

As explained in section 2, the income and mileage pair \((y_i, m_i)\) varies across ZIP Code areas, so the heterogeneity within each ZIP Code area exclusively comes from \(\epsilon_{ijt}\) in equation (15). Therefore, the consumer surplus of each new car buyer in the same ZIP Code area is

\[
CS_{it} = \frac{1}{\beta_i} \ln[1 + \sum_{j=1}^{J} \exp(\delta_{jt} + \mu_{ij})].
\] (20)

When implementing alternative taxes, the household-specific utility component becomes

\[
\mu_{ij}^f = \beta \frac{p_j}{y_i} + \beta \lambda \frac{q + \tau_f}{y_i} e_j m_i
\]

and

\[
\mu_{ij}^g = \beta \frac{p_j + e_j \tau_g}{y_i} + \beta \lambda \frac{q}{y_i} e_j m_i
\]

accordingly. Given these formulas, I compute the change in consumer surplus to measure
gains and losses for consumers as

$$\Delta CS_i^k = CS_i - CS_i^k$$

$$= \frac{\ln(1 + \sum_{j=1}^{J} \exp(\delta_{jt} + \mu_{ij})) - \ln(1 + \sum_{j=1}^{J} \exp(\delta_{jt} + \mu_{ij}^k))}{\beta_i},$$

in which $k \in \{f, g\}$ stands for the specific type of a new tax.

Appendix A.2 shows

$$\frac{\partial CS_{it}^f}{\partial \tau^f} = \frac{\beta \lambda}{\beta} m_i \sum_{j=1}^{J} s_{ijt} e_j$$

and

$$\frac{\partial CS_{it}^g}{\partial \tau^g} = \sum_{j=1}^{J} s_{ijt} e_j.$$

It appears that the change in consumer surplus caused by applying a new fuel tax is associated with vehicle usage. In contrast, a tax on low fuel economy affects the consumer surplus in a more uniform pattern. Figure 1 (b) mirrors the differentiated distributive implications of a fuel tax and a tax on low fuel economy reflected in the previous figure. It shows that the loss in total consumer surplus associated with the implementation of a fuel tax is smaller relative to that of a tax on low fuel economy. This result is consistent with those calculated in Grigolon et al. (2018), in which they suggest that a lower tax amount is required to achieve the same externality reduction when mileage heterogeneity is properly accounted for.

6. Conclusion

This paper has shown that fuel taxes are less regressive than fuel economy taxation policies when both are designed to increase the market share of HEV/BEV by the same amount in the new car market. I account for the price effect of policies by estimating a demand model in which new car buyers are allowed to re-optimize their choices after implementing alternative taxes. Specifically, I pair a random coefficients logit model of
new car demand with a joint distribution of household income and vehicle usage. The heterogeneity allowed in such a framework unmasks differentiated distributional impacts rooted in the targeting mechanism of policies.

This result helps shed light on designing equitable policy instruments for achieving the transition toward a more sustainable transport system. This paper suggests that the overall welfare impact of taxes and subsidies is associated with the elasticity of demand for driving, which should not be overlooked. It is also possible to project results produced in this paper to markets of other energy-using durable goods. When both energy taxes and energy efficiency standards are feasible for generating energy savings, properly accounting for usage heterogeneity is crucial in evaluating alternative policies on distributional grounds.
Online Appendix

Appendix A  Theoretical appendix  29
  Appendix A.1  The income elasticity of tax revenue  29
  Appendix A.2  Change in consumer surplus  32

Appendix B  Data appendix  33
  Appendix B.1  Alternative to gasoline-powered cars  33
  Appendix B.2  MAVC sample  34

Appendix C  Empirical appendix  35
  Appendix C.1  Additional results from policy counterfactuals  35
Appendix A. Theoretical appendix

Appendix A.1. The income elasticity of tax revenue

Let

\[ \tilde{\phi}_{ijt} = x_j^I \alpha + \beta (\tilde{p}_j + \lambda \tilde{q}_e_j m_i) / y_i + \zeta_t + \zeta_{jt}, \quad (A.1) \]

in which \( m_i \) is a function of \( y_i \). This yields

\[ \frac{\partial \tilde{\phi}_{ijt}}{\partial y_i} = -\frac{\beta \tilde{p}_j}{y_i^2} + \frac{\beta \lambda \tilde{q}_e_j \partial m_i}{y_i} - \frac{\beta \lambda \tilde{q}_e_j m_i}{y_i^2} > 0 \quad \text{if} \quad \frac{y_i m_i \partial m_i}{\partial y_i} < 1. \quad (A.2) \]

Given

\[ \tilde{s}_{ijt} = \frac{\exp(\tilde{\phi}_{ijt})}{1 + \sum_{j=1}^{J} \exp(\tilde{\phi}_{ijt})}, \quad (A.3) \]

we have

\[ \frac{\partial \tilde{s}_{ijt}}{\partial y_i} = \frac{\exp(\tilde{\phi}_{ijt}) \frac{\partial \tilde{\phi}_{ijt}}{\partial y_i}}{1 + \sum_{j=1}^{J} \exp(\tilde{\phi}_{ijt})} + \exp(\tilde{\phi}_{ijt}) \frac{\sum_{j=1}^{J} \exp(\tilde{\phi}_{ijt}) \frac{\partial \tilde{s}_{ijt}}{\partial y_i}}{[1 + \sum_{j=1}^{J} \exp(\tilde{\phi}_{ijt})]^2} \quad (A.4) \]

\[ = \tilde{s}_{ijt} \frac{\partial \tilde{\phi}_{ijt}}{\partial y_i} + \tilde{s}_{ijt} \sum_{j=1}^{J} \tilde{s}_{ijt} \frac{\partial \tilde{\phi}_{ijt}}{\partial y_i} > 0. \quad (A.5) \]

Because

\[ T_{it}^{f} = m_i \tau^{f} \sum_{j=1}^{J} \tilde{s}_{ijt} e_j \quad (A.6) \]

and

\[ T_{it}^{g} = \tau^{g} \sum_{j=1}^{J} \tilde{s}_{ijt} e_j, \quad (A.7) \]
then

\[
\eta_f^y = \frac{\partial T^f_{it}}{\partial y_i} y_i
\]

(A.8)

\[
= (m_i \tau^f \sum_{j=1}^{J} \frac{\partial \tilde{s}_{ijt}}{\partial y_i} e_j + \frac{\partial m_i}{\partial y_i} \tau^f \sum_{j=1}^{J} \tilde{s}_{ijt} e_j) \frac{y_i}{m_i \tau^f \sum_{j=1}^{J} \tilde{s}_{ijt} e_j}
\]

(A.9)

\[
= y_i \frac{\sum_{j=1}^{J} \frac{\partial \tilde{s}_{ijt}}{\partial y_i} e_j}{\sum_{j=1}^{J} \tilde{s}_{ijt} e_j} + y_i \frac{\partial m_i}{\partial y_i} \frac{\sum_{j=1}^{J} \tilde{s}_{ijt} e_j}{m_i \tau^f \sum_{j=1}^{J} \tilde{s}_{ijt} e_j}
\]

(A.10)

\[
= y_i \frac{\sum_{j=1}^{J} \frac{\partial \tilde{s}_{ijt}}{\partial y_i} e_j}{\sum_{j=1}^{J} \tilde{s}_{ijt} e_j} + \eta^m_y
\]

(A.11)

and

\[
\eta^g_y = \frac{\partial T^g_{it}}{\partial y_i} y_i
\]

(A.12)

\[
= (\tau^g \sum_{j=1}^{J} \frac{\partial \tilde{s}_{ijt}}{\partial y_i} e_j) \frac{y_i}{\tau^g \sum_{j=1}^{J} \tilde{s}_{ijt} e_j}
\]

(A.13)

\[
= y_i \frac{\sum_{j=1}^{J} \frac{\partial \tilde{s}_{ijt}}{\partial y_i} e_j}{\sum_{j=1}^{J} \tilde{s}_{ijt} e_j}.
\]

(A.14)

This derivation gives

\[
\eta^f_y = \eta^g_y + \eta^m_y,
\]

(A.15)

in which \(\eta^f_y\) is the income elasticity of the fuel tax revenue, \(\eta^g_y\) is the income elasticity of the tax revenue for a tax on low fuel economy, and \(\eta^m_y\) is the income elasticity of driving.

Although equation (1) uses a static setup, following Allcott and Wozny (2014), it’s possible to assume a forward-looking new car buyer will divide this \(h_{ij}\) into two parts: the present discounted value of fuel costs during her holding period, and the present discounted value of fuel costs over the remainder of the car’s life after it is resold. In this model, a car owner knows the payout of selling this car to its next owner will be the present discounted value of the resale price plus that present discounted value of the remaining fuel cost. When
considering the present discounted value of the fuel tax revenue, equation (6) becomes

\[ \hat{T}_{it} = \lambda m_i \tau^f \sum_{j=1}^{J} \tilde{s}_{ij} e_j. \]  

(A.16)

Similarly, the income elasticity of the present discounted value of the fuel tax revenue is

\[ \hat{\eta}_y = \frac{\partial \hat{T}_{it}}{\partial y_i} \frac{y_i}{\hat{T}_{it}} \]  

(A.17)

\[ = (\lambda m_i \tau^f \sum_{j=1}^{J} \frac{\partial \tilde{s}_{ij} e_j}{\partial y_i} + \lambda \frac{\partial m_i}{\partial y_i} \tau^f \sum_{j=1}^{J} \tilde{s}_{ij} e_j) \frac{y_i}{\lambda m_i \tau^f \sum_{j=1}^{J} \tilde{s}_{ij} e_j} \]  

(A.18)

\[ = y_i \sum_{j=1}^{J} \frac{\partial \tilde{s}_{ij} e_j}{\partial y_i} + y_i \frac{\partial m_i}{m_i \frac{\partial y_i}{\partial y_i}} \]  

(A.19)

\[ = y_i \sum_{j=1}^{J} \frac{\partial \tilde{s}_{ij} e_j}{\partial y_i} + \eta^m \]  

(A.20)

\[ = \eta^g + \eta^m. \]  

(A.21)
### Appendix A.2. Change in consumer surplus

Let

\[ CS_{it} = \frac{1}{\beta_i} \ln[1 + \sum_{j=1}^{J} \exp(\delta_{jt} + \tilde{\mu}_{ij})], \quad (A.22) \]

in which

\[ \tilde{\mu}_{ij} = \beta \frac{p_j + e_j \tau^g}{y_i} + \beta \lambda \frac{q + \tau^f}{y_i} e_j m_i. \quad (A.23) \]

Then

\[ \frac{\partial CS_{it}}{\partial \tau^f} = \frac{1}{\beta_i} \frac{1}{1 + \sum_{j=1}^{J} \exp(\delta_{jt} + \tilde{\mu}_{ij})} \frac{\beta \lambda}{m_i} \sum_{j=1}^{J} \exp(\delta_{jt} + \tilde{\mu}_{ij}) e_j \quad (A.24) \]

\[ = \frac{\beta \lambda}{\beta} m_i \sum_{j=1}^{J} s_{ijt} e_j, \quad (A.25) \]

and

\[ \frac{\partial CS_{it}}{\partial \tau^g} = \frac{1}{\beta_i} \frac{1}{1 + \sum_{j=1}^{J} \exp(\delta_{jt} + \tilde{\mu}_{ij})} \beta \sum_{j=1}^{J} \exp(\delta_{jt} + \tilde{\mu}_{ij}) e_j \quad (A.26) \]

\[ = \sum_{j=1}^{J} s_{ijt} e_j \quad (A.27) \]
Appendix B. Data appendix

Appendix B.1. Alternative to gasoline-powered cars

Table B.4: Non-Fossil Fuel Alternatives

<table>
<thead>
<tr>
<th>Model</th>
<th>Model Year 2011</th>
<th>Model Year 2012</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Battery Electric Vehicle</strong></td>
<td></td>
<td></td>
<td>47</td>
</tr>
<tr>
<td><strong>Flexible Fuel and Hybrid Electric Vehicle</strong></td>
<td></td>
<td></td>
<td>753</td>
</tr>
</tbody>
</table>

Table B.4: According to this WEB PAGE, Tesla Model S began deliveries to US customers in June of 2012. In December 2012, Tesla Motors was granted its first a full Class 1 Dealer License in Massachusetts.
### Table B.5: New Vehicle Registration Counts in 2011

<table>
<thead>
<tr>
<th>Model Year</th>
<th>Car</th>
<th>SUV</th>
<th>Truck</th>
<th>Van</th>
<th>Car</th>
<th>SUV</th>
<th>Truck</th>
<th>Van</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010 and earlier</td>
<td>1,078</td>
<td>178</td>
<td>106</td>
<td>77</td>
<td>64</td>
<td>17</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>2011</td>
<td>7,788</td>
<td>4,882</td>
<td>1,120</td>
<td>549</td>
<td>5,119</td>
<td>3,488</td>
<td>1,091</td>
<td>405</td>
</tr>
<tr>
<td>2012</td>
<td>364</td>
<td>34</td>
<td>0</td>
<td>28</td>
<td>4,631</td>
<td>2,131</td>
<td>274</td>
<td>275</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>9,230</td>
<td>5,094</td>
<td>1,226</td>
<td>654</td>
<td>9,814</td>
<td>5,636</td>
<td>1,372</td>
<td>682</td>
</tr>
</tbody>
</table>

Table B.5: When comparing to the count of registered vehicles published by the Massachusetts Department of Revenue (DOR), which is the only other official public source of vehicle registration counts, the vehicle census was 25% and 32% lower than the DOR estimate for year 2009 and 2010. The Metropolitan Area Planning Council Data Services Department suggests users to use these data with caution. So a complete enumeration of registered vehicles in the MAVC starts since 2011. A vehicle is marked as newly purchased if its first vehicle history record is also the first owner history record, and the starting odometer reading of this record is smaller than 300 miles. Following Reardon et al. (2016), several criteria are employed to flag and remove low quality observations in the MAVC. About 70% of the vehicles included in the demand sample have three consecutive inspection records from 2011 through 2014. These observations are applied for extracting annual mileage at the ZIP Code level.
Appendix C. Empirical appendix

Appendix C.1. Additional results from policy counterfactuals

Table C.6: Simulated New Fuel Tax

<table>
<thead>
<tr>
<th>HEV/BEV Market Share Increase</th>
<th>New Fuel Tax ($/gallon)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>0.2598</td>
</tr>
<tr>
<td>2%</td>
<td>0.5013</td>
</tr>
<tr>
<td>3%</td>
<td>0.7281</td>
</tr>
<tr>
<td>4%</td>
<td>0.9426</td>
</tr>
<tr>
<td>5%</td>
<td>1.1468</td>
</tr>
<tr>
<td>6%</td>
<td>1.3423</td>
</tr>
<tr>
<td>7%</td>
<td>1.5305</td>
</tr>
<tr>
<td>8%</td>
<td>1.7122</td>
</tr>
<tr>
<td>9%</td>
<td>1.8885</td>
</tr>
<tr>
<td>10%</td>
<td>2.0599</td>
</tr>
</tbody>
</table>

Table C.7: Gallon-per-mile Tax for a 1% Expansion of HEV/BEV

<table>
<thead>
<tr>
<th>Car Fuel Economy Rating</th>
<th>the tax on low fuel economy at Purchase ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 MPG</td>
<td>5,728</td>
</tr>
<tr>
<td>25 MPG</td>
<td>4,582</td>
</tr>
<tr>
<td>30 MPG</td>
<td>3,819</td>
</tr>
<tr>
<td>35 MPG</td>
<td>3,273</td>
</tr>
<tr>
<td>40 MPG</td>
<td>2,864</td>
</tr>
</tbody>
</table>

Table C.7: Allcott et al. (2014) have proposed an Internality Targeting Principle for deriving the optimal policy to address consumer undervaluation of energy cost. In their automobile market simulations, the second best policy is a gasoline tax combined with a product subsidy that would decrease the purchase price of 25 vs. 20 MPG vehicles by about $700, which is comparable to the subsidy scheme obtained from policy counterfactuals carried out in this paper.
References


