Title of Presentation: Consumer Heterogeneity and Gasoline Price Response: Implications for Optimal Tax policy

By

Edson Okwelum

Graduate Student

Department of Environmental and Natural Resource Economics

University of Rhode Island

Email:Edson_okwelum@my.uri.edu

Title of Presentation: Consumer Heterogeneity and Gasoline Price Response: Implications for Optimal Tax policy

Edson Okwelum

Department of Environmental and Natural Resource Economics

University of Rhode Island

November 30, 2014

Abstract

Measuring consumer response to gasoline price changes is a fundamental issue in the design and regulation of environmental externalities. In this paper, we document the importance of accounting for heterogeneity in consumer utilization of durable goods in explaining the apparent undervaluation of future fuel costs. We develop a Bayesian method within the context of heterogeneous discrete choice model paired with pricing equations derived from Bertrand competition to estimate heterogeneous demand elasticity for gasoline price changes, and use our results to conduct counterfactual analyses of alternative tax policies. We find that accounting for heterogeneity in utilization and other dimensions all but eliminates undervaluation of future operating costs. Results from our counterfactual analyses imply that gasoline taxes lead to welfare increases that are 20% higher than those obtained under a fuel economy regime.

Keywords: *Bayesian Econometrics, Heterogeneity, Gasoline Prices, Gasoline Policy, Simulation*

1. Introduction

In the last few decades, the U.S. government has introduced several taxes and subsidies on the purchase of new vehicles that depend on the performance of vehicle fuel economy. The main driver for taxation of fuel economy is to correct externalities that result from fuel consumption. The main regulatory tools available to regulators are several and result in different outcomes. Gasoline taxes are generally efficient and also are the very easy administer. In addition, gasoline taxes also directly target the source of the externality. However, for political reasons, they politicians do not like to increase the price of gasoline (Stevens, 2006). However, if the driving behavior of households are inelastic and consumers show undervaluation of the future operating costs savings when making purchase decisions, then a tax scheme that targets directly the vehicle choice behavior of households would equally be efficient at tackling the externalities in automobiles (Grigolon et al 2014; Allcott and Greenstone (2012)). A product tax such as fuel economy taxation on the other hand, are most favored by regulators and are less efficient than gasoline taxes. Several arguments have been given for the prevalent of fuel economy taxation (Sallee 2010).

This paper develops an empirical approach for estimating equilibrium demand and supply in differentiated products markets and then applies these techniques to analyze relative efficiency of gasoline based policies while accounting for consumer heterogeneity. This paper imagines a distribution of consumer preferences over products which are subsequently paired with pricing equations derived from Bertrand competition to obtain parameter estimates under equilibrium conditions[1]. The estimation approach is Bayesian, in a heterogeneous demand and supply framework model. We estimate the model using a Hierarchical Bayesian framework and use a Monte Carlo Markov Chain (MCMC) method with data augmentation to simulate the posterior distribution of the parameters. We favor this approach for several reasons. First, we are concerned that product characteristics that are observed to the consumer, but not by the analyst are correlated with the error terms. Therefore, we make the identifying assumptions that the unobserved product characteristics are not independent of observed product characteristics.

To compare different gasoline policies, we model the household choice of miles; vehicle attributes (size), and other goods. We assume that consumers differ by income and two taste parameters- miles and vehicle size. We are concerned mainly with the relative efficiency of the

---

[1] This is important for counterfactual policy simulations.

policy instruments in the presence of heterogeneous choices about gasoline and car characteristics and not with the distributional effects (Bento et al 2012; Sandmo 1975). We use dat data from the 2000 San Francisco Bay Area travel survey (BATS). The main data set is from a 2-day survey of 15000 households and contains information on vehicle fleet mix of households, individual and household socio-demographics, individual characteristics and activities.

In this paper, we contribute to several lines of research. First, we contribute to the literature on *Energy Paradox*[2] in fuel economy by providing empirical evidence that unobserved consumer heterogeneity could result in different consumers sorting into different vehicle types based on their valuation of fuel economy. This implies that the coefficient on fuel economy also reflects values on other attributes that are associated with fuel economy. Unlike existing literature in this topic, this paper accounts for sorting bias due unobserved heterogeneity by using an "equilibrium sorting"[3] model that uses a mixture of distributions to characterize unobserved and observed heterogeneity among households (Train and Winston 2007). The "equilibrium sorting" model uses the properties of market equilibria, as well as information on the behavior of economic agents, to infer structural parameters that characterize agent heterogeneity. With the random utility model, we allow annual vehicle miles travelled, time preferences and expectations of gasoline prices to vary across consumers. We treat the discounted operating costs and vehicle costs as random variables. This allows us to obtain a distribution of households' preferences for fuel economy across the population.

Recent works on how consumers value fuel economy tend to use monthly within vehicle variation over time in gasoline prices to identify consumer tradeoffs between fuel costs and vehicle costs. The question is that in the presence of sorting bias due to observed and unobserved consumer heterogeneity, whether this is an accurate way to identify tradeoffs between vehicle costs and fuel economy. Panel data and individual fixed effects provide potential solutions when very large data sets spanning long period are available. This source of identification would not be sufficient to identify consumer weighting of future fuel costs in our case because vehicle prices would be considered to be endogenous. Additionally, firms respond to gasoline prices in the

---

[2] Energy Paradox is defined as the disconnect between net present value estimates of i energy conserving cost savings and what consumers actually pay on energy conservation (Metcalf and Hasset 1999, Jaffe et al 2001).
[3] The paper uses a straight forward extension of the framework used by Bayer and Timmins (2003) to examine sorting in

short term by adjusting vehicle costs to match sales. A key aspect of this is that prices are negotiated by dealers, and will depend upon inventories. For example, if gas guzzlers are not selling, dealers will offer price discounts. When demand is high for vehicles with high mileage, discounts are lower, and dealers could charge a premium over the retail price. For example, MacManus (2005) finds that the shift to higher fuel efficiency vehicles brought about by rising gasoline prices were obscured by price cuts disproportionately aimed at gas guzzlers. In addition, we also account for volatility in gasoline price by allowing households to continually update their expectations of gasoline price. In future time periods, households dynamically adjust their driving habits in response to gasoline price.

The consumer heterogeneity we find is substantial and significant. The heterogeneity arises from huge differences in the amount of miles travelled by consumers and heterogeneity in expectation of gasoline prices and time preferences. We find that a substantial portion (61%) of upper the 95% of households in our sample correctly value fuel economy as they are willing to pay $0.99 to reduce obtain a $1.00 discounted future gasoline costs over the lifetime of the vehicle. And 29% of the upper 95% overvalue fuel economy as they are willing to pay an average of $1.57.

How consumers weigh temporal effects of future fuel costs have important policy implications. This is because the nature of the temporal weighting helps in determining if market failure exists and helps indicate whether policy prescriptions that affect initial vehicle costs such as gas guzzlers tax will reduce fuel consumption at lower costs than gasoline tax. In addition, results from such "equilibrium sorting" models can be used to develop theoretically consistent predictions for the welfare implications of future policy changes on fuel economy and gasoline taxes. Results from our counterfactual analyses imply that gasoline taxes are 20% more efficient than to fuel economy and leads welfare increases.

We contribute to several lines of literature. First are papers comparing different policy instruments targeting automobile externality (Fullerton and West 2002, Sandmo 1975, 1976). However, this paper is different from those because we use a microcosmic famework. This work is closely related to the work of Bento et al (2012), Sawhill (2008), Alcott and Wozny (2009) Bayer *et al* (2011). However, this paper extends the literature in several respects. Our work is different from Sawhill's in several respects[4]. First, we use individual household data why he

---

[4] Both papers use random coefficient logit models.

used aggregate data. In addition, while we both control for price endogeneity (Sawhill uses BLP's contraction mapping method and does not include information about pricing behavior). His does not include a supply model. The main difference between this work and these papers is on the identification strategy and the estimation of the supply side model. First unlike Bento et al (2012), this work is an empirical study rather than a simulation and such provides empirical evidence of the effect of consumer heterogeneity in estimates of the consumer trade-off of fuel economy and vehicle costs.

This paper is structured as follows. Section two discusses the equilibrium model composed of a demand side and supply side model. The section three, present the data used for the analysis. Section Four discusses the identification and estimation strategy as well as the issues when estimating the specification. This is followed by the policy simulation and finally, we conclude in six.

2. Methodology and Conceptual Framework

The model is consistent with a structural equilibrium model of a heterogeneous product competition. The approach is based on earlier models by Yang et al (2003), Berry et al (1995) and Petrin (2002) and but with significant differences. We simultaneously estimate models of demand and supply with household heterogeneity. The estimation strategy is divided into two different steps. In the first step, the paper estimates household-level demand functions and then aggregates these individual functions to construct estimated firm demand curves. In the second step, the paper then uses estimated demand curves to solve firms' first order conditions under the assumptions of Bertrand –Nash competition. The demand side of the model is based on a random utility function of consumer vehicle choice following Berry, Levinsohn and Pakes (1995; henceforth, BLP).

a. Consumer Preferences for fuel economy

The demand for a new vehicle is viewed as an intertemporal choice problem in which consumers' trade-off future fuel savings and vehicle price. Each household derives utility from both vehicle ownership and utilization. Consumers' choice of vehicle type is specified as a random parameter logistic model similar to Allenby and Lenk (1994), Chintaninga et al (2001), Berry et al (1995) and Villas-Boas and Winer (1999). We assume that consumers gain utility from vehicle miles and size and other goods.

A Household i, who maximizes utility from choosing vehicle j or not (the outside good is used to capture utility other than new cars) in each choice occasions $t=1,....T$ solves the following optimization problem in equation 1.

$$\max_{j \in \{0,...,J\}} u_{ijt} = \alpha_i\left(y_{it} - p_{jt} + \gamma_i G_{ijt}\right) + x_{jkt}\beta_k + \xi_{jt} + \varepsilon_{ijt} \tag{1}$$

where u_{ijt} is consumer I's utility for model j in time t. x_{jkt} represent k^{th} vector of observed product and individual characteristics interacted with demographic variables including brand intercepts. By specifying the x_{ij} appropriately, you can construct any pattern of covariance across alternatives. y_{it} is consumer i's income while ξ_{jt} represent unobserved product attributes[5]. G_{ijt} is the expectation on gasoline price changes for consumer i. The terms in ξ_{jt} represent unobserved utility component that induces correlation as well as substitution in vehicles to overcome the independent of irrelevant alternative problem (IIA) (Train 2003). ε_{ijt} is unobserved error term that is assumed to be correlated with automobile prices. The error term is iid distributed across consumers with Type 1 extreme value. We introduce heterogeneity between subjects by allowing some in the intercept and some characteristics have random variation. The heterogeneity arises from consumers having different taste for size and mileage. If z_{mij} is the subset of covariates with random variation, and coefficient $\beta_{m,}$ such that for household $i,$ the heterogeneity model have both a systematic and random component itself:

$$\beta_{im} = \beta_m + \eta_m D_i + \varepsilon_{im} \qquad (2a)$$
$$\gamma_{im} = \gamma_m + \eta_m D_i + \varepsilon_{im} \qquad (2b)$$

The D_i represents household related characteristics that could help in explaining the heterogeneity. α_i, β_i and γ_i are individual specific taste parameters. We allow preferences over price to vary with income so that α_i represents the individual specific marginal utility of income. We make the following assumptions about the error terms and individual specific taste parameters:

$$\theta = (\alpha_i, \beta_i')' \sim MVN(\hat{\theta}, \Sigma)^6 \qquad (3a)$$
$$\xi_{jt} \sim MVN(0, \Sigma) \qquad (3b)$$

$$\varepsilon_{ijt} \sim iid\ N(0,1) \qquad\qquad (3c)$$

Given the type 1 extreme value distribution of the idiosyncratic error term, the probability that household i choose alternative j on choice occasion t, conditional on the model structural parameters is given by

$$\Pr{}_{it}(j\,|\,\xi) = \frac{e^{\alpha_i(y_{it}-p_{jt}+\gamma_i G_{ijt})+x_{jkt}\beta_k+\xi_{jt}}}{\sum e^{\alpha_i(y_{it}-p_{jt}+\gamma_i G_{ijt})+x_{jkt}\beta_k+\xi_{jt}}} \qquad\qquad (4a)$$

The value of ξ is unobserved for each individual, although we can estimate it by drawing from a known density function. To obtain the unconditional choice probability for each individual, the logit probability is integrated over all values ξ weighted by its density function:

$$P_j = \int\int_\xi \left[\frac{e^{\alpha_i(y_{it}-p_{jt}+\gamma_i G_{ijt})+x_{jkt}\beta_k+\xi_{jt}}}{\sum e^{\alpha_i(y_{it}-p_{jt}+\gamma_i G_{ijt})+x_{jkt}\beta_k+\xi_{jt}}} \right] dF_\xi(\xi;\theta) \qquad\qquad (4b)$$

b. Supply Model and Equilibrium Prices

This section develops the supply side of the model in which vehicle manufacturers set prices to maximize profits, given price of its competitors. Vehicles are differentiated, that is two very similar vehicles from different manufacturers will be priced differently. We assume that firms, $f,= 1,.....F,$ compete in a Bertrand Nash fashion under a differentiated product. The manufacturer sets prices $p_j= (p_{1j,......}p_{Tj})$. In the short term, firms only change prices. Firm f's profit function is given by:

$$\Pi_{f,t} = \sum_{j\notin J_{f,t}} \left(p_{jt} - c_{jt}\right) S_j N \qquad\qquad (5)$$

where p_j is as defined earlier, N is the number of US households, S_j is the predicted market share obtained by summing individual consumers weighted sum of vehicle selection probabilities, c is the unit variable cost of product j. Following the procedure developed by Villas-Boas (2007)[7], we solve equation 4 for the first order conditions with respect to price.

$$\frac{\partial \Pi_{ft}}{\partial p_{jt}} = S_j + \sum_j (p_j - c_j)\frac{\partial S_{jt}}{\partial p_j} = 0 \quad (6)$$

---

[7] Also Villas-Boas and Winer 1999

The first order condition can be summarized in matrix form by the following equation with number of rows equal to the number of models in the market.

$$\mathbf{mc} = \mathbf{p} - \nabla(\mathbf{p}, \mathbf{X}; \xi)^{-1} \mathbf{S}(\mathbf{p}, \mathbf{X}; \xi) \qquad (7)$$

where ∇ is a J x J matrix whose terms are given by:

$$\Delta_{j\eta} = \begin{cases} -\dfrac{\partial S_j}{\partial p_j} & \text{if } j \text{ and } r \text{ are produced by same firm} \\ 0 & \text{otherwise} \end{cases}$$

From the first order conditions in a Bertrand Game, automobile prices depend on the marginal costs of a product. Solving for the first order conditions, one obtains equations that satisfy the price-cost mark-ups $p_j - mc_j$[8] and market share for each product that satisfies the function given below. The vector of mark-ups only depends on the parameters of the demand equations and equilibrium price vector. We can easily use 6 for counterfactual policy simulation to solve for new equilibrium price vectors.

For a firm that produces model j in period t, he marginal cost depends on both observed product attributes and unobserved product attributes:

$$mc_{jt} = \omega(X_{jt,}c_{jt}, \eta) + \mu_j \qquad (8)$$

where $\omega(.)$ is a parameter function. μ_j is the unobserved idiosyncratic cost associated with model j. X_{jt} may include all the same characteristics that affect demand.

Using 5 and 7 and with a little algebra, we can specify the each manufacturer's cost as model specific linear function of cost shifters, ω:

$$p_{jt} = \omega(X_j, c_j, \eta) + mc_{jt} + \frac{\partial S_j}{\partial p_j} S + \mu_j \qquad (9)$$

μ is the supply side error term that is distributed Multivariate normal and is also correlated with price. As such, the demand and supply side error terms are correlated with price through the demand side error term.

$$\mu \sim MVN(0, \Sigma_s) \qquad (10)$$

3. **Estimation Strategy**

---

[8] The Betrand Model without product differentiation implies price equals marginal cost. Because we are assuming product differentiation, this is not the case here.

We use a simultaneous equations model limited information approach for estimation. This is mainly driven by the fact that there is a correlation between price and demand equation error terms. As such we need to use instruments. In addition, price is also correlated with the error term in the cost equation. Because of the correlation of the price with the error terms in both the demand and supply equations, we use a simultaneous-equations technique to gain efficiency (Sudhir 2001). Therefore, we need an instrumental variable-based Bayesian Simultaneous equations estimation approach. The covariance matrix between the error terms of the demand and supply equations has the following distributions:

$$\begin{pmatrix} \xi_{jt} \\ \mu_{jt} \end{pmatrix} \sim N(0, \Sigma); \ \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \tag{11}$$

Because of price endogeneity, the covariance between σ_{12} error terms in the two equations results will not be equal to zero. By controlling for this correlation with a simultaneous estimation procedure, we control for any unobserved variables that may bias the coefficient of price in the demand equation. We assume that, z, is the exogenous and independent of the ξ, μ terms and as such we have $E(\xi z) = 0$, and $E(\xi z) = 0$ as well. That is we assume that z is correlated with price, but not the unobservable affecting the outcome variables. Under weak additional assumptions such as assuming a bivariate normal distribution of the unobservables; and the effect of the covariates on the price is given by cdf of a standard normal distribution evaluated at the at a point is additive to in the *X's* and *z,* we can identify the parameter coefficients given instruments. The implication of the model structure is that the marginal density of the outcome variable is not normal because this would ignore the dependence between the error terms. Rather we will use the joint choice-price models to derive the form of the conditional density function and by expressing the errors in terms of the elements of the covariance matrix[9]. In this case, the model reduces to a standard simultaneous equation model except that we have latent endogenous variables. Just as is obtainable in simultaneous equations, we require exclusion restrictions to aid with identification.

Identification issues in this specification mirror those that are encountered in classical simultaneous demand and supply equations. The identification follows ideas developed by Li (1993); Heckman (1978) and Chib (203). While the coefficients in each equation are not identifiable from the separate equations, identification is achieved by setting one of ($\sigma11$, $\sigma22$)

---

[9] Chib (2003) provides an exhaustive treatment of the approach

=1 (Drez and Richard 1983). This is because the coefficient of the instruments is identified given the marginal distribution of the X's. Because of the importance of the instruments for identification argument, we need an idea of the correlation between instruments with the endogenous variable after controlling for the effects of the other covariates. As part of the sampling strategy, we obtain the posterior distribution of the

We estimate the model using a Hierarchical Bayesian framework and use a Monte Carlo Markov Chain (MCMC) method with data augmentation to simulate the posterior distribution of the parameters. We discuss the MCMC in the Appendix. For potential instrumental variables, we follow BLP (1995) and use as instruments, the sum of the characteristic across own-brand products (excluding that product), and the sum of the characteristic across rival firm products.

4. Data

a. Data

The data sets used for the analysis were obtained from several sources. The primary data used for this analysis is drawn from the 2000 San Francisco Bay Area travel survey (BATS). The main data set is from a 2-day survey of 15000 households and contains information on vehicle fleet mix of households, individual and household socio-demographics, individual characteristics and activities. The sample used for this analysis, however includes 3500 households with either 1, 2, or 3 vehicles. The vehicles owned by each household are categorized into one of 16vehicle types based on their vintage and size.

The dataset contains detailed information on all cars owned by households surveyed as well as information on family size, income and several other demographic attributes. Most importantly, the dataset includes the mileage of each car owned by the household during each quarter. Table 1 provides information on demographics[10] of the estimation sample, and it is consistent with U.S population's socioeconomic data for California. This table also shows significant variations in household characteristics across the vehicle classes. Wealthier households (as measured by total yearly expenditures) also possess larger vehicles. Also,

---

[10] Among the wide range of household demographic information the survey contains includes size of household, age of household members, sex, employment status, type of residence, whether or not the household is located in a rural or urban area, education level, household income

households with more workers or income earners and those with male heads are inclined to have SUVs.

The information on automobiles includes the make, model, vintage and MSRP of each car, and a large set of vehicle characteristics. I will supplement these datasets with data on vehicle characteristics from EPA fuel economy test data and Automotive News Market Data Book. The former provides information on fuel economy measured in miles per gallon[11]. While the latter includes information on size, performance, and standard options of various models. Data on annual observations of certain microeconomic and macroeconomic variables also included as well as gasoline prices. Summary statics of the attributes of the estimation model is provided in Table 2.

Information on gasoline prices (incl. state and local taxes) is taken from the U.S. Energy Information Administration (EIA). EIA collects weekly retail gasoline prices for all formulations (conventional, oxygenated and reformulated gasoline prices), which are recorded by State. We match individual households in the automobile purchase dataset to state-level gasoline prices based on the respondents state of origin. Finally, we also collect information on interest rates on new and old car loans from the Federal Reserve and Consumer Price Index for all goods from the Bureau of Labor Statistics. Next, we will describe how the paper handles consumer's expectations of gasoline prices.

> b. *Discounted Future Gasoline Price Expectations*

Gasoline prices are generally difficult to predict. When consumers make vehicle purchase decisions, how they form expectations for future gasoline prices is not well known. The most common assumption past works have made is that consumers treat the gasoline price process as a random walk, with any price shock considered permanent with subsequent adjustments in demand (Biesebroeck and leuven 2010) . However if the price shocks decay rapidly, then estimated elasticity may be biased because price shock will have less of an effect on demand and measured price elasticities will be lower.

Consistent with prior literature, we will make the assumption that current gasoline prices are best predictors of the expected future gasoline prices, i.e. gasoline prices are follow a random walk process with time trend. Support for the random walk hypothesis comes from Anderson *et*

---

[11] The mpg values are combined EPA mpg values: a weighted average of City and Highway MPG values that is calculated by weighting the City value by 55% and the Highway value by 45%.

al (2012) who concluded that the random walk process is a good reflection of consumer expectations of future gasoline prices. Additional support is provided by the result of a dickey fuller test in which we fail to reject the null hypothesis that average monthly gasoline prices exhibits a unit root. Figure 1 is a time series plot of the natural log of the real price of gasoline prices between January 1994 to December 2012. In addition, we also try a gasoline price expectations model in which we use only lagged data to generate sets of forecasts that use information from the customers to update their most recent forecasts. Figure 1b is a plot of the forecasts from the model using this procedure superimposed over forecast that just uses one period ahead.

> *Parameter Estimates of Demand Model*

We use proper but relatively diffuse priors for the specification. Estimates of the posterior means and their standard deviations are presented in and standard errors from the specifications are presented in Tables 4. The two variables of interest are capital cost of vehicle (α) and discounted annual operating costs (γ)[12]. Given that there is no specific direction in existing literature for other variables to include; we include weight, horsepower, size (wheelbase and track width), drive train, dummies if vehicle is an SUV, minivan, full sized van or pickup truck, and indicators that reflect make of a given vehicle. The indicators capture extent of household preference for a particular make of vehicle. The demographic variables enter through interactions with vehicle attributes. The parameters vector has a multivariate normal distribution. We divide the coefficient on price by the income of consumer to allow the elasticity to vary with income.

Column 1 of Table 4 reports the posterior means for the parameters while Colum 2 reports the standard deviations. The coefficient on price and operating cost are negative as expected. The mean price coefficient is 3.58 and significant. The standard deviation of a normally distributed price coefficient is 0.50.[13] We recover mean elasticity with respect to discounted operating costs of 14.39 with a t-value of 14.44. We also find standard deviation of

---

[12] Fuel economy enters the consumers' utility function as annual operating costs. This is determined as real price of gasoline per gallon divided by mpg of fuel multiplied by the average annual vehicle miles. This formulation accounts for the wide variations in the price of gasoline in different locations and in annual miles travelled by different households. Importantly, it implies an inverse relationship between increasing fuel economy and utility since fuel economy has an inverse relationship with fuel consumption, while fuel expenditure have a linear relationship with consumption.

[13] This implies that 95% of the price coefficient over all models varies between 2.6 to 4.56.

the normally distributed coefficient of discounted yearly operating costs of 1.05 which implies that 95% of the drivers have mean elasticities ranging from 5.57 to 23.21.

We also report ratios of the coefficients of the two variables of interest: $\frac{\gamma}{\alpha}$ with a mean value of 1.82. A Wald test of the null hypothesis that the coefficients on price and discounted operating costs are equal fails to reject the hypothesis at 5% level (W=2.48<χ^2<3.65). We find that 95% of the consumers in our sample have $\frac{\gamma}{\alpha}$ values ranging from 1.65 to 6.0. This large spread signifies heterogeneity in consumer response to fuel economy. In this vein, we uncover three segments of discounted operating costs: 29% of the consumers significantly overvalue fuel economy while 61% less so or are rational. We profile the segments in terms of their average demographics characteristics. On average, rational households are lager and tend to have above average income, have higher number of school age children, men older than 45 years. The correlations matrix shows that owners of large vehicles have errors that are negatively correlated with those of smaller vehicles. This reflects the contrasting effects of the two main predictors for these two groups. It also shows us that incomes effects outweigh slightly outweigh household size for new car owners, but the reverse is the case for used cars.

c. Do Consumers Undervalue Fuel Economy

Table 5presents estimates of several ratios of $\frac{\gamma}{\alpha}$ computed over a range of discount rates and vehicle lifetime values. These values can be interpreted as the consumer willingness to pay for reduced future discounted costs. Expressing the results in this form makes it easy and the interpretation is intuitive. However, these ratios can be easily translated into a form that allows the values to be comparable to results reported elsewhere. We achieve this by dividing the ratio $\frac{\gamma}{\alpha}$ by$\sum_{s=0}^{T}\left(\frac{1}{1+r}\right)$. Using data from NHTSA vehicle survivability and travel mileage schedules (NHTSA 2006), we assume the average vehicle has a useful life (T) of 20 years. To check sensitivity, we try different values for T 10, 15, 20 and 25. Also, we assume discount rates ranging from 5-9%. Most published work have used discount rates or either 6 or 7%. The values in Column 3 are comparable to those reported by Alcott and Wozny (2009) who used a discount rate of 6% for their base model.

The last two columns of Table 5 report our results. We estimate that for every $1dollar saved in the future, consumers take no more than $0.99 into account. We find evidence of small

overvaluation and not one instance of undervaluation. These results are in stark contrast to results reported by Helfand and Wolverton 2010; Greene 2010; Alcott and Wozny 2009. We find that accounting for heterogeneity and sorting by consumers, tends to remove any significant evidence of consumer undervaluation of future fuel savings. The last two rows present lower 5% and upper 95% distribution of consumer valuation of fuel economy. Though the lower 5% consumers only place values of $0.50 of a dollar savings, it does not signify energy paradox by this population in the strict sense.

5. Policy Counterfactual

a. Tax Policies under Consumer Heterogeneity

In this section, we use the consumer utility form described earlier to model how consumers respond to different fuel economy polices. The two main policies we look at are the gasoline tax and tax on the efficiency of vehicles (fuel economy). The gasoline tax is a tax on gasoline consumption and if set at the marginal abatement costs, it is very efficient[14]. The fuel economy is a product tax that is equivalent to a tax on the vehicle's CO_2 emissions. In the simplest form, a gasoline tax regime induces households to drive fewer miles, buy fuel-efficient vehicles (Fullerton and West 2002) and smaller vehicles (Auffhammer 2013). Even though different households would respond in different ways to the gasoline policy, it would induce the same behavior, Pigou (1932). However, the reality is that consumers demand for mileage, vehicle attributes and other factors affect the outcome of these policies in diverse ways. Second, emissions are not invariant to automobile design which implies that there are other factors that affect emissions levels.

In this section, I wish to focus on the economic efficiency of different gasoline tax instruments in the presence of consumer heterogeneity while abstracting from the distributional implications of such policies[15]. To understand how consumer heterogeneity affects the efficiency of different policy instruments, we incorporate them into the indirect utility function in 1:

$$u_{ijt} = \alpha_i \left(\left(y_{it} - \left(p_{jt} + t_j^E \right) \right) + \gamma_i \frac{\left(g + t_j^g \right)}{mpg\left(size_i \right)} * miles \right) + x_{jkt} \beta_k + \xi_{jt} + \varepsilon_{ijt} \qquad (12)$$

The household problem under 12 is to maximize utility with respect to miles, size and m. The optimal gasoline tax, product tax as well as size based tax could be obtained from solving the simultaneous equations from the first order conditions of the utility maximization of problem 12 for each household. Previous work by Fullerton and West (2003) and Innes (1996) have looked at such problems. Salle (2010); Allcott and Greenstone (2012) also provides a discussion of the merits of the different gasoline policies. The gasoline tax induces households to drive the optimal miles even with heterogeneous taste parameters and it cannot vary with different consumers' choice of vehicle. The fuel economy taxation on the hand targets the different vehicle classes differently. Evidence from previous research shows that fuel economy taxation does have an impact of fleet fuel economy through its influence on market shares of targeted vehicles (Sallee 2010).

I use 12 to simulate the impact of different gasoline tax policies that target either gasoline consumption or fuel economy. Our simulation is driven by the following assumptions. First, the durable goods nature of automobiles implies that we have to account for both consumers' taste for fuel efficiency and intensity of utilization in estimating elasticity. Therefore, households make a two-step decision. They first choose the number and type of vehicles to own based on driving habits and expectations of future gasoline prices. And in the second step, conditional on their vehicle bundles, household choose utilization intensity which determines fuel consumption. The second decision influences decisions about vehicle type choice in the first step. Second, the error term of the different dimension vehicle choice and utilization are correlated. This correlation could be in either direction. For example, if the unobserved taste parameter which induces some households to choose fuel efficient vehicles also induces them to drive less, then the error terms will be positively correlated. We can also imagine a situation in which the unobserved taste parameters induce heavy drivers to increase utilization, resulting in the error terms being negatively correlated.

Also, about 38% of the households in our sample have two vehicles and 7% owning multiple vehicles. This lends to "portfolio effects" as a possible interpretation for households owning more than one vehicle and how they respond. That is the mix of vehicles owned by households satisfies different functions. We follow Dube(2004) and Hendel(1999) in accounting for households ownership of portfolio of vehicles in a mixed logit with repeated choice occasions framework. Because we do not observe each choice occasion in our data set, we are unable to

precisely characterize the actual situations in which households consume each alternative. However, we do observe the mile driven by each vehicle type, and this allows us to model the distribution of consumption occasions and make inference about the context in which they are consumed. One can relate different choice occasions as representing different activities undertaken by the household that requires driving such as school runs, shopping, weekend family gateways, etc. Intuitively, household vehicle portfolio and annual mileage indicates heterogeneity in tastes for various choice occasions.

b. Effect of Different Policies on Market Shares

Table 6 reports the results of our counterfactual analysis of the effects of gasoline tax and fuel economy tax on market shares. We simulate the effect of both a gasoline tax and fuel economy tax on each household in our sample and then aggregate them by quartile in terms of fuel economy to obtain the effect. In California, the average gasoline tax in California is 63.8c/gallon (this includes 18c federal gasoline tax). A 25c rise in the gasoline tax has the effect of increasing the market share of the most efficient vehicles in our model by about 0.2 points; while product tax results in increases in the market share of the most efficient category quartile of vehicles in our sample by 0.13 points. At the other extreme, a gasoline tax will decrease the market share of the lease efficient vehicles in our sample by 1.87 points; while a product tax reduces the market share of the least efficient group of vehicles in our sample by 1.56 points. Taken in together, a gasoline tax is approximately about 20 percent more effective than a product tax in the whole fleet.

6. Conclusions

In this paper, I present empirical evidence of how consumer heterogeneity affects the relative effectiveness of different gasoline policy instruments. Using a Bayesian approach that is within the context of heterogeneous discrete choice model paired with pricing equations derived from Bertrand competition to estimate heterogeneous demand elasticity for gasoline price changes, we first provide evidence of the effect of consumer heterogeneity in obtaining estimates of how consumers value fuel economy. Thus, consumer heterogeneity matters for whether the first best outcome could be or for that matter how to set the second-best rates. We then build a

model of heterogeneous consumers that differ by income, tastes for miles, and tastes for engine size. Then we use the parameter estimates to conduct a counterfactual simulation analysis.

We find no evidence to support the argument that consumers systematically underweight the cost of future events in real market settings. However, we find significant evidence that different consumers sort into different groups such that rational consumers accurately tradeoff future operating costs and vehicle prices. This research has several limitations which provide a useful roadmap for future studies. There is no evidence that consumers value new vehicles and old vehicles in a similar version or that different consumers drive their vehicles over their useful life as developed in this paper. In terms of policy, we find those consumers' mileage heterogeneity results in gasoline taxes being 20% more efficient than fuel economy taxes. In fact, evidence from textbook market exists that forward and myopic consumers place different values on the resale price of durable goods. The policy results should be viewed with caution because the dataset we use is specific to California alone. Issues may arise when attempts are made to translate the results to other states especially if the driving habits of residents of California are markedly different from other states. Ideally, the model could be extended to allow for dynamic vehicle replacement as well creating a fully dynamic framework. Second, the model developed in this paper makes strong assumptions about the functional form of the distribution of consumers in the population for ease of estimation. By restricting consumer taste parameters for price and operating costs to be normally distributed, we are necessarily ruling out more flexible distributions of consumer behavior which may be present in the population.

References

Ackerberg, Daniel A., and Marc Rysman. "Unobserved Product Differentiation in Discrete-Choice Models: Estimating Price Elasticities and Welfare Effects." *RAND Journal of Economics* 36, no. 4 (2005): 771–88.

Allcott, Hunt. "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market." *American Economic Journal: Economic Policy* 5, no. 3 (August 2013): 30–66. doi:10.1257/pol.5.3.30.

Allcott, Hunt, and Nathan Wozny. *Gasoline Prices, Fuel Economy, and the Energy Paradox*. Working Paper. National Bureau of Economic Research, November 2012. http://www.nber.org/papers/w18583.

Amil Petrin and Kyoo-il Kim, Control Function Corrections for Omitted Attributes in Differentiated Product Models, 2010, Working Paper.

Anderson, Soren T., Ryan Kellogg, and James M. Sallee. "What Do Consumers Believe about Future Gasoline Prices?" *Journal of Environmental Economics and Management* 66, no. 3 (2013): 383–403.

Athey, Susan, and Guido W. Imbens. "Discrete Choice Models with Multiple Unobserved Choice Characteristics*." *International Economic Review* 48, no. 4 (November 1, 2007): 1159–92. doi:10.1111/j.1468-2354.2007.00458.x.

Bento, Antonio M., Shanjun Li, and Kevin Roth. "Is There an Energy Paradox in Fuel Economy? A Note on the Role of Consumer Heterogeneity and Sorting Bias." *Economics Letters* 115, no. 1 (2012): 44–48.

Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile Prices in Market Equilibrium." *Econometrica* 63, no. 4 (1995): 841–90.

Berry, Steven T. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25, no. 2 (1994): 242–62.

Brownstone, David, David S. Bunch, and Kenneth Train. "Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-Fuel Vehicles." *Transportation Research Part B: Methodological* 34, no. 5 (June 2000): 315–38. doi:10.1016/S0191-2615(99)00031-4.

Busse, Meghan R, Christopher R Knittel, and Florian Zettelmeyer. "Are Consumers Myopic? Evidence from New and Used Car Purchases." *American Economic Review* 103, no. 1 (February 2013): 220–56. doi:10.1257/aer.103.1.220.

Chinn, Menzie D., and Olivier Coibion. "The Predictive Content of Commodity Futures." *Journal of Futures Markets* 34, no. 7 (July 1, 2014): 607–36. doi:10.1002/fut.21615.

DeLeire, Thomas, Shakeeb Khan, and Christopher Timmins. "Roy Model Sorting and Nonrandom Selection in the Valuation of a Statistical Life*." *International Economic Review* 54, no. 1 (February 1, 2013): 279–306. doi:10.1111/j.1468-2354.2012.00733.x.

Dreyfus, Mark K., and W. Kip Viscusi. "Rates of Time Preference and Consumer Valuations of Automobile Safety and Fuel Efficiency." *Journal of Law and Economics* 38, no. 1 (1995): 79–105.

Gillingham, K. Forthcoming. Identifying the Elasticity of Driving: Evidence from a Gasoline Price Shock in California. *Regional Science and Urban Economics*

Goldberg, Pinelopi Koujianou. "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry." *Econometrica* 63, no. 4 (1995): 891–951.

Greene, D. 2010. How Consumers Value Fuel Economy: A Literature Review. Technical Report EPA-420-R-10-008.

Hamilton, James D. *Understanding Crude Oil Prices*. NBER Working Paper. National Bureau of Economic Research, Inc, 2008. http://ideas.repec.org/p/nbr/nberwo/14492.html.

Hassett, Kevin A., and Gilbert E. Metcalf. "Energy Conservation Investment: Do Consumers Discount the Future Correctly?" *Energy Policy* 21, no. 6 (June 1993): 710–16. doi:10.1016/0301-4215(93)90294-P.

Hausman, Jerry A. "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables." *Bell Journal of Economics* 10, no. 1 (1979): 33–54.

Hughes, Jonathan E., Christopher R. Knittel, and Daniel Sperling. "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand." *The Energy Journal* 29, no. 1 (2008): 113–34.

Kahn, James A. "Gasoline Prices and the Used Automobile Market: A Rational Expectations Asset Price Approach." *The Quarterly Journal of Economics* 101, no. 2 (May 1, 1986): 323–39. doi:10.2307/1891118.

Klier, Thomas, and Joshua Linn. "The Price of Gasoline and New Vehicle Fuel Economy: Evidence from Monthly Sales Data." *American Economic Journal: Economic Policy* 2, no. 3 (August 2010): 134–53. doi:10.1257/pol.2.3.134.

Mannering, Fred, and Clifford Winston. "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization." *RAND Journal of Economics* 16, no. 2 (1985): 215–36.

McManus, Walter. *The Effects of Higher Gasoline Prices on U.S. Light Vehicle Sales, Prices, and Variable Profit by Segment and Manufacturer Group, 2001 and 2004*. Industrial Organization. EconWPA, 2005. http://ideas.repec.org/p/wpa/wuwpio/0505010.html.

———. *The Link between Gasoline Prices and Vehicle Sales: economic Theory Trumps Conventional Detroit Wisdom*. MPRA Paper. University Library of Munich, Germany, 2007. http://ideas.repec.org/p/pra/mprapa/3463.html.

Parry, Ian W. H, Margaret Walls, and Winston Harrington. "Automobile Externalities and Policies." *Journal of Economic Literature* 45, no. 2 (June 2007): 373–99. doi:10.1257/jel.45.2.373.

Petrin, Amil. *Quantifying the Benefits of New Products: The Case of the Minivan*. NBER Working Paper. National Bureau of Economic Research, Inc, 2001. http://ideas.repec.org/p/nbr/nberwo/8227.html.

Petrin, Amil, and Kenneth Train. *Control Function Corrections for Unobserved Factors in Differentiated Product Models*. 2006 Meeting Paper. Society for Economic Dynamics, 2006. http://ideas.repec.org/p/red/sed006/856.html.

Sallee, James M., and Joel Slemrod. "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." *Journal of Public Economics* 96, no. 11 (2012): 981–99.

Sallee, J.M., S. E. West and W. Fan, 2010. "The Effect of Gasoline Prices on the Demand for Fuel Economy in Used Vehicles: Empirical Evidence and Policy Implications", Draft version 1.1 of January 19, 2010.

Sawhill, J.W., 2008. Are Capital and Operating Costs Weighted Equally in Durable Goods Purchases? Evidence from the U.S. Automobile Market, discussion paper, Department of Economics, University of California at Berkeley, Berkeley, California, April.

Train, Kenneth. *Discrete Choice Methods with Simulation*. Online economics textbooks. SUNY-Oswego, Department of Economics, 2003. http://ideas.repec.org/b/oet/tbooks/emetr2.html.

Train, Kenneth E., and Clifford Winston. *Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers*. Working paper. Regulation2point0, 2006. http://ideas.repec.org/p/reg/wpaper/331.html.

Van Biesebroeck, Johannes. *The Demand for and the Supply of Fuel Efficiency in Models of Industrial Organisation*. OECD/ITF Joint Transport Research Centre Discussion Papers. Paris: Organisation for Economic Co-operation and Development, January 1, 2010. http://www.oecd-ilibrary.org/content/workingpaper/5kmjp66c9vjc-en.

Verboven, Frank. "Quality-Based Price Discrimination and Tax Incidence: Evidence from Gasoline and Diesel Cars." *The RAND Journal of Economics* 33, no. 2 (2002): 275. doi:10.2307/3087434.

Villas-Boas, J. Miguel, and Russell S. Winer. "Endogeneity in Brand Choice Models." *Management Science* 45, no. 10 (October 1, 1999): 1324–38. doi:10.1287/mnsc.45.10.1324.

Villas-Boas(2007) A Note on Limited versus Full Information Estimation in Non-Linear Models, Working Paper, Haas School of Business, University of California, Berkeley.

Table 1: Description of Sample Households who Bought New Vehicles

Socioeconomic Characteristics	
Variable	Sample Value
Average Household Income	$68,908
Average Age	50
Average Household Size	2.61
Percentage Male	57
Percentage with Child 1-6	15

Table 2: Summary Statistics of Sample Vehicle Characteristics

Automobile Characteristics		
Variable	Mean Value	Standard Deviation
Annual Vehicle Miles Travelled (Miles)	11,861	17930
Real Gasoline Price ($/gallon)	1.75	0.51
Vehicle Price	$22,716	$10,448.05
Length (Inches)	192.1	16.5
Wheelbase (Inches)	112.9	11.4
HP (Pounds)	194	52
Curb Weight (Pound)	3479	695
Mpg	20.16	1.45

Table 3: ARIMA (1,1,1) Model Annual US Retail Gas Price Data 1994-2005

Variable	Estimate	Standard Error
Log of Real Gasoline Price		
Constant	0.5208	0.126
ARMA		
AR1	0.8593	0.0691
MA1	0.4671	0.1182
Variance of Residuals (σ^2)	0.0102	0.0074

Table 4: Parameter Estimates of Consumer Demand Model

Demand Side Parameters	Mean (1)	SD (2)	2.5% (3)	97.5%
Posterior Mean(β's)				
Constant	-8.093	1.009	-11.250	-6.093
Vehicle Cost/Income ($'000)-α	-3.58	-1.08	-5.36	-0.48
Operating Costs/Income-ϒ	-6.78	2.085	-6.78	-0.01
Vehicle Size (Wheelbase * Length)	3.460	1.38	2.89	3.98
Log HP	9.736	0.143	-10.007	29.552
Curb Weight (*tons*)	18.931	13.65	-7.82	45.68
Log of HH Size	0.18	0.21	0.36	1.83
Log of Engine Size	4.169	5.19	-19.126	27.46
r11	0.240		-0.155	0.928
r21	-0.297			
r31	-0.0820		-0.496	0.003
r41	-1.817		-4.659	1.0253
r22	-0.4592			
r32	-0.1851			
r42	-0.2664			
r33	-0.0014			
r43	-0.0289		-0.325	0.267
r44	-0.0101			
ϒ/α	1.84			
# Observations	3,460			

Table 5: Robustness of Consumer Valuation of Fuel Economy

	Ratio Υ/α=1.88		
	(1)	(2)	(3)
	T=10	T=15	T=25
Discount Rate=6%	1.540	1.170	0.890
Discount Rate=7%	1.620	1.250	0.980
Discount Rate=8%	1.770	1.460	1.150
Discount Rate=9%	1.690	1.330	1.060
Upper 95%	0.99-1.88		
Lower 5%	0.78		

Table 6a: Effects of Gasoline and Fuel Tax on Market Shares

	Market Share	Change in Market Share	
		25c Gasoline Tax	Product Tax
Fuel Economy Q1- Least Efficient	14.7	-1.65	-1.56
Fuel Economy Q2	21.1	-0.40	-0.35
Fuel Economy Q3	28.1	0.20	0.27
Fuel Economy Q4- Most Efficient	36.8	0.19	0.13