

July 2012 ■ RFF DP 12-30

# The Heterogeneous Effects of Gasoline Taxes

*Why Where We Live Matters*

---

**Elisheba Spiller and Heather M. Stephens**

1616 P St. NW  
Washington, DC 20036  
202-328-5000 [www.rff.org](http://www.rff.org)



# **The Heterogeneous Effects of Gasoline Taxes: Why Where We Live Matters**

Elisheba Spiller and Heather M. Stephens

## **Abstract**

Using disaggregated confidential household data, we estimate spatial variation in household-level gasoline price elasticities and the welfare effects of gasoline taxes. A novel approach allows us to model a discrete-continuous household choice of vehicle bundles, while disaggregating the choice set and including vehicle-specific fixed effects and unobserved consumer heterogeneity. The mean elasticity of demand for gasoline is  $-0.67$ , but with tremendous variation across location and income. We find that rural households have 30 percent more negative welfare impacts than urban households from gasoline taxes. Finally, we explore different policies that can help to mitigate welfare inequalities due to these taxes.

**Key Words:** gasoline taxes, welfare, elasticity, rural, commuting, transportation

**JEL Classification Numbers:** Q0, R0, H0

© 2012 Resources for the Future. All rights reserved. No portion of this paper may be reproduced without permission of the authors.

Discussion papers are research materials circulated by their authors for purposes of information and discussion. They have not necessarily undergone formal peer review.

## **Contents**

<b>I. Introduction .....</b>	<b>1</b>
<b>2. Model.....</b>	<b>4</b>
<b>3. Estimation .....</b>	<b>10</b>
<b>4. Data .....</b>	<b>13</b>
<b>5. Results .....</b>	<b>17</b>
<b>6. Welfare Analysis &amp; Policy.....</b>	<b>24</b>
<b>7. Conclusion .....</b>	<b>31</b>
<b>References .....</b>	<b>33</b>

# The Heterogeneous Effects of Gasoline Taxes: Why Where We Live Matters

Elisheba Spiller and Heather M. Stephens\*

## 1. Introduction

Gasoline consumption by U.S. households accounts for at least 17 percent of total greenhouse gas emissions in the United States (U.S. EPA [2011]). Additionally, about 49 percent of all petroleum consumed in the United States comes from foreign sources (U.S. EIA [2011]). Given the threats of air pollution, climate change, and political instability in countries that provide foreign oil, policy makers have struggled with how to reduce gasoline consumption. Economists tend to focus on gasoline taxes as an optimal solution to the externalities associated with driving, with the added benefit of providing revenue for the government. However, understanding the disaggregated effects of increases in gasoline prices requires considering how demand elasticity and welfare effects vary across households.

In general, gasoline taxes can reduce the amount of vehicle miles traveled and gasoline consumed, though households vary in their ability to reduce their use of gasoline and mitigate the negative welfare impacts of higher gasoline prices. For example, individuals with access to public transit can commute to work by subway when gasoline prices increase- Goldberg (1998) finds that those in the Northeast with greater access to public transit drive less. Furthermore, there is evidence that rural households drive much more than urban households (Schmalensee and Stoker [1999], Gillingham [2011]), though the size of the urban area is important, as large metro areas can result in increased driving patterns (West [2004]). As household location affects the ability of households to adjust their consumption of gasoline when faced with higher gasoline prices, it implies that gasoline taxes will have diverse welfare effects depending on home location.

Vehicle ownership also affects elasticities of demand for gasoline: households with more than one vehicle can substitute driving to the more efficient vehicles when faced with high gasoline prices (Berkowitz et al. [1990], Feng, Fullerton and Gan [2005], Spiller [2011]). As variations in household characteristics, location, and vehicle ownership results in diverse

---

\* Elisheba Spiller, Postdoctoral Fellow, Resources for the Future, Washington, DC, [spiller@rff.org](mailto:spiller@rff.org). Heather M. Stephens, Assistant Professor, Department of Economics, California State University, Long Beach, [heather.stephens@csulb.edu](mailto:heather.stephens@csulb.edu)

elasticity estimates and driving patterns across households, it is important to account for these differences when estimating the distributional effects of gasoline taxes.

This is especially important given the wide range in demographics and income across different regional areas. For example, Table 1 shows the means in our dataset of different household demographics across regional areas depending on their level of urbanization (with the first column being the most urban). This Table shows that urban households are richer (though not relative to housing prices), have more fuel efficient vehicles, own less vehicles, commute more, and have more children relative to adults. This variance in demographics and vehicle ownership across different regional areas highlights the importance of accounting for household location when analyzing the welfare impacts of gasoline prices.

While there has been extensive research demonstrating the need to account for regional variation in economic outcomes when conducting policy analysis (e.g. Ferguson et al. [2007], Partridge et al. [2008a], and Porter [2003]), the elasticity literature has mostly failed to account for regional heterogeneity in gasoline price sensitivities. In fact, most have estimated national-level elasticities (Klier and Linn [2010]), or have simply grouped people by demographics and broad regional groups. Although a few papers have considered rural and urban households separately (e.g. Bento et al. [2008], Gillingham [2011], and Schmalensee and Stoker [1999]) this still masks the extreme diversity within both rural and urban areas. Partridge, Ali, and Olfert (2010) found that the ability of rural households to integrate into urban areas through commuting depends on their distance to urban centers, demonstrating the importance of allowing for heterogeneity in welfare impacts of gasoline taxes even within different regions. Therefore, our paper improves the elasticity literature by incorporating significant geographic heterogeneity (within and across different regional areas) into the estimation of welfare effects due to gasoline taxes.

Our model allows us to calculate the elasticity of demand for gasoline for each household in our dataset, accounting for both how much people drive (the intensive margin) and what types of vehicles they choose to purchase (the extensive margin). We use detailed, geographically defined data to also account for how far an individual lives from a metropolitan area, as well as the local cost of living at the county level. Utilizing confidential, household-level data from the 2009 National Household Travel Survey (NHTS) we estimate the elasticities and welfare effects associated with gasoline taxes.

**Table 1. Demographics by Urban Specification**

Beale* Code	1	2	3	4	5	6	7	8	9
Income	623,098	544,281	528,812	464,203	512,914	447,719	425,671	436,527	397,746
Income/ median rent	0.61	0.71	0.74	0.69	0.72	0.73	0.70	0.70	0.67
Vehicle miles travelled	18,259	17,623	17,548	17,532	16,990	19,054	16,999	20,625	16,783
MPG	22.12	21.90	21.68	21.31	21.45	20.66	20.91	20.69	20.83
# Vehicles	1.78	1.76	1.81	1.78	1.75	1.85	1.79	1.87	1.83
Commute time	14.64	10.68	9.98	9.06	7.33	9.64	7.36	11.51	8.17
Commute distance	10.53	8.38	8.02	7.95	6.00	8.74	6.62	10.95	6.50
# Adults	1.97	1.97	1.97	1.98	1.98	1.97	1.96	2.00	2.01
Household size	2.27	2.15	2.14	2.05	2.06	2.06	2.01	2.12	2.07

\* Beale codes defined below:

- 1 County in metro area with 1 million population or more
- 2 County in metro area of 250,000 to 1 million population
- 3 County in metro area of fewer than 250,000 population
- 4 Nonmetro county with urban population of 20,000 or more, adjacent to a metro area
- 5 Nonmetro county with urban population of 20,000 or more, not adjacent to a metro area
- 6 Nonmetro county with urban population of 2,500-19,999, adjacent to a metro area
- 7 Nonmetro county with urban population of 2,500-19,999, not adjacent to a metro area
- 8 Nonmetro county completely rural or less than 2,500 urban population, adj. to metro area
- 9 Nonmetro county completely rural or less than 2,500 urban population, not adj. to metro area

Furthermore, we utilize a novel method of estimating elasticities of demand for gasoline that allows the researcher to eliminate many assumptions on household behavior that are usually employed in the literature. These modeling assumptions restrict the ability of the researcher to capture more subtle changes in consumer behavior, which can cause an underestimation of the elasticity of demand for gasoline, and thus impact the welfare estimates of a gasoline tax. This estimation technique was first introduced in Spiller's (2011) working paper and is extended here to allow for richer geographic heterogeneity at the household level.

We find significant variation across households in terms of sensitivity to gasoline prices. The elasticity estimates vary widely depending on household characteristics and location. For example, richer households and households with more than one vehicle are much more sensitive to gasoline prices than are poorer households with only one vehicle. We also find substantial heterogeneity in terms of location effects, where both elasticities and welfare impacts vary widely depending on distance to urban centers and regional location. Finally, we propose a revenue recycling mechanism that the government can provide to mitigate the welfare inequities associated with a gasoline tax while still accruing revenue.

The rest of the paper is organized as follows: Section 2 outlines the model; Section 3 provides details on our estimation approach; Section 4 describes the data; Section 5 presents the elasticity results; Section 6 analyzes the welfare impacts of a gasoline tax and proposes revenue recycling policies; and Section 7 concludes.

## 2. Model

Our model is more flexible than most models utilized in the elasticity literature, and allows the researcher to reduce the number of restrictive assumptions that are commonly placed on the household's decision making process for ease of estimation. This flexibility decreases the chance that we will underestimate how elastic households truly are, providing us with a more precise estimate of elasticity and thus welfare effects of gasoline taxes. We are able to employ a more flexible model by utilizing a novel method that has not been used in the literature and is described in more detail in Section 4. In this section, we demonstrate how the different flexibilities we incorporate into our model can minimize the underestimation of the elasticity of demand for gasoline.

We model the household as a unified decision maker, so that each household optimally chooses the number and types of vehicles to own and how many miles to drive each of the vehicles in its garage. This implies that each household maximizes utility such that it optimally allocates vehicle miles traveled (VMT) among the vehicles according to characteristics such as fuel efficiency (and thus relative operating costs), size, horsepower, and comfort. Furthermore, the choice of vehicles and VMT depends on the household's characteristics. For example, if a household is very large, they may prefer to drive larger vehicles in order to accommodate a larger number of individuals at any time. Another important factor that affects vehicle choice is location. For instance, in an urban center with scarce parking, households may prefer smaller vehicles. Away from an urban center, households may prefer more fuel efficient vehicles in order to minimize expenses associated with traveling to work or to urban amenities. However, rural

households also may prefer trucks or 4WD vehicles if they live on a farm or in an area with unpaved roads or inclement weather. Thus, household characteristics will affect the types of vehicles chosen, and so we specifically model this interaction between household and vehicle characteristics.

Similarly, the amount of VMT allocated to each vehicle in the garage will also depend on the relative characteristics of *all* the vehicles owned by the household. Our model allows households to change their behavior in terms of which vehicles are used for specific purposes. For example, if a household owns both a car and an SUV, we may see the household allocating a larger VMT to the car when faced with high gasoline prices. On the other hand, if gasoline prices are low, a household that has a high commute time may choose to allocate VMT to the SUV in order to benefit from the higher level of comfort that this vehicle provides.

By allowing households to optimize over a bundle of vehicles, we also minimize the downward bias on the gasoline demand elasticity estimate. Not allowing households to optimize over their entire bundle of vehicles can cause households to appear much less elastic than they truly are.<sup>1</sup> This is due to the fact that the household's ability to shift VMT among the vehicles that they own is another margin along which the household can adjust when faced with higher gasoline prices. If the researcher doesn't allow the household to shift along this third margin, she or he may miscalculate the household's ability to respond to gasoline prices, thus underestimating the elasticity of demand for gasoline. Therefore, by modeling the household's decisions over its bundle of vehicles, we minimize this underestimation. Furthermore, we find that households with more than one vehicle are substantially more responsive to gasoline prices, thus supporting the claim that not allowing for substitution among vehicles in the household's garage could lead to underestimation of the elasticity.

To model household-level demand for gasoline, we follow the utility specification as in Bento et al. (2003), adapted as in Spiller (2011) to allow for the optimization across the household's bundle of vehicles. The household's indirect utility function is defined in equation (1).

---

<sup>1</sup> See Spiller (2011) for a discussion of the impacts of the independence assumption on the elasticity estimate.

$$V_i = \frac{-1}{\lambda_i} \exp\left(-\lambda_i \left(y_i - \sum_{k=1}^{J_i} P_{ik}^u\right)\right) - \frac{1}{\sum_{k=1}^{J_i} \beta_{ik}} \exp\left(\alpha_i + \sum_{k=1}^{J_i} \beta_{ik} P_{ik}^d\right) + \sum_{k=1}^{J_i} (\theta_k * \theta_s + \tau_{ik} + \varepsilon_{ik}) \quad (1)$$

Here  $i$  indexes household while  $j$  indexes vehicle.  $J_i$  is the total number of vehicles household  $i$  owns. The coefficients  $\alpha$ ,  $\beta$ , and  $\tau$  are interacted with vehicle ( $z_j$ ) and household ( $z_i$ ) characteristics (where  $z_{ij} = z_i * z_j$ ) so that different households can have different preferences for vehicle characteristics, although these are not random utility parameters and are thus not household specific. On the other hand,  $\lambda$  is household specific and allows households to have unobserved preferences for driving; this is estimated non-parametrically. These parameters are defined as:

$$\begin{aligned} \alpha_i &= \alpha z_i^\alpha \\ \lambda_i &= f(z_i^\lambda) \\ \beta_{ij} &= -\exp(\beta z_{ij}^\beta) \\ \tau_{ij} &= \tau z_{ij}^\tau \\ \varepsilon_{ij} &\sim N(0, \sigma^2) \end{aligned}$$

$P_{ij}^u$  is the used vehicle price of vehicle  $j$  for household  $i$  and is calculated using a hedonic price gradient based on vehicle characteristics and state-month gasoline prices. To account for financing possibilities that would minimize overall yearly payments, we assume that the household pays only 70 percent of the used vehicle price.<sup>2</sup>  $P_{ij}^d$  is the operating cost of vehicle  $j$  for household  $i$  and is equal to the price of gasoline in household  $i$ 's state in the month the interview was conducted, divided by the vehicle's EPA adjusted miles per gallon.

$\theta_j$  is the fixed effect of vehicle  $j$ , which captures all the characteristics, such as comfort or style, that are unobserved by the econometrician but observed by both the consumer and producer. As Berry et al. (1995) proved, not including a vehicle-specific fixed effect can create a bias on the vehicle price coefficient. If a vehicle has a strong unobserved vehicle characteristic,

<sup>2</sup> Utilizing a discount in yearly payments is common in hedonics literature, see Bishop and Timmins (2011), page 10 for a discussion. We discount less than they do, but this is due to a shorter payback period on vehicle loans compared to mortgages.

such as being very stylish, the producer can charge more for that vehicle, and, all else being equal, consumers will likely purchase it in greater quantities. Thus, if we do not account for this unobserved characteristic, it will appear that consumers have a preference for expensive vehicles rather than stylish vehicles, causing an upward bias on the price coefficient. The exclusion of vehicle-specific fixed effects can also cause a distortion on the gasoline demand elasticity estimate. In this case, if a more fuel efficient vehicle has a negative unobserved quality and is thus purchased in smaller quantities, it would appear that individuals care less about fuel efficiency than they actually do, and thus the elasticity could be underestimated. While this bias could go either way (i.e. individuals could instead appear to care more about fuel efficiency than they actually do), Spiller (2011) finds that the elasticity is in fact underestimated substantially when the researcher does not include vehicle-specific fixed effects. Therefore, we include vehicle-specific fixed effects into the model in order to minimize the bias on both the price coefficient and elasticity estimate.

In order to better identify substitution patterns across different types of vehicles, we define a vehicle as a model-year-nameplate, such as a 1996 Honda Civic, instead of aggregating across vehicle types into larger groups such as car, SUV, van, or truck. Aggregating different vehicles into larger groups can cause individuals to appear less responsive to gasoline prices, as it masks more subtle movements that consumers can make along the extensive margin.<sup>3</sup> By disaggregating the choice set into model-year-nameplates, we are able to capture more of this movement and thus not underestimate the elasticity of demand for gasoline.

$\theta_s$  is a state-fixed effect. Here we interact vehicle-fixed effects with state-fixed effects in order to capture preference differences across vehicles in different geographic locations. For example, there may be a stronger preference for unobserved style in California, while in the Midwest there may be a stronger preference for unobserved comfort. By allowing the vehicle-fixed effect to vary by state, we do not force equal preferences for unobserved vehicle attributes across geographic locations. This allows for greater consumer heterogeneity in the model.

Through Roy's Identity, we can find the optimal VMT for each vehicle (as defined in equation (2)). Thus our use of an indirect utility function accounts for the optimal choice of VMT in an integrated utility framework.

---

<sup>3</sup> See Spiller (2011) for impacts on elasticity given aggregation of choice sets.

$$VMT_{ij}^* = \frac{\beta_{ij}}{\sum_{k=1}^{J_i} \beta_{ik}} \exp \left\{ \sum_{k=1}^{J_i} (\alpha_{ik} + \beta_{ik} P_{ik}^u) + \lambda_i \left( y_i - \sum_{k=1}^{J_i} P_{ik}^d \right) \right\} \quad (2)$$

Equation (2) implies that the optimal VMT for each vehicle depends not only on its own characteristics but also on the characteristics and fuel efficiency of the other vehicles that the household owns. In this way, if the *relative* operating cost of the household's second vehicle increases due to higher gasoline prices, then the optimal VMT of the household's first vehicle will increase. For example, if a household owns both an SUV and a car, then an increase in gasoline prices would result in a shift in VMT from the SUV to the car. This substitutability provides the household more flexibility when dealing with higher gasoline prices, and not allowing for this can cause the household to appear less elastic than it truly is.

The household characteristics that we include are household size, log of household income, ratio of household income to median rent, average commute time, ratio of number of household drivers to number of vehicles, distance from the household's census tract to the center of the closest metropolitan statistical area (MSA), census tract population density, and a dummy variable equal to one if the household lives in a rural area.<sup>4</sup> The vehicle characteristics we include are MPG, size (length\*width), acceleration (horsepower/weight), a truck dummy, vehicle age, and vehicle length. The full set of interactions between  $z_i$  and  $z_j$  are listed in Table 2. (Each row is interacted, so, for example, distance to MSA is interacted with the truck dummy in  $z_{ij}^\beta$ .)

The interactions in  $\beta$  and  $\tau$  allow preferences of vehicle characteristics to vary based on household characteristics. Thus, if, for example, the parameter on household income\*vehicle age is negative, this implies that richer households prefer newer vehicles. While the magnitude of the parameters is hard to interpret in these sorts of nonlinear equations, the sign and significance describes consumer preferences. Berry, Levinsohn, and Pakes (2004) emphasized the importance of including interactions between household and product characteristics: "Models without individual differences in preferences for characteristics generate demand substitution patterns that are known to be a priori unreasonable (depending only on market shares and not on the characteristics of the vehicles)" (p. 72). Thus, we interact household characteristics with vehicle characteristics in order to better understand substitution patterns in vehicle choice. The variables

---

<sup>4</sup> Further explanation of the data is contained in Section 5.

in our model are those based on available data from NHTS that we believe are most likely to predict driving patterns and vehicle ownership.<sup>5</sup>

**Table 2. Interaction Parameters**

Parameter	$z_i$	$z_j$
$z_i^\alpha$	• Household size	
	• Household income/median rent	-
	• Commute time	
	• Number of drivers/number of vehicles	
	• Rural area dummy	
$z_{ij}^\beta$	• Commute distance	• MPG
	• Household size	• Size
	• Household income/median rent	• Acceleration (HP/weight)
	• Distance to MSA	• Truck dummy
$z_{ij}^\tau$	• Household income	• Vehicle age
	• Census tract population density	• Length

$\lambda_i$  is a function of household characteristics that accounts for unobserved preferences and varies across households. If equation (2) did not contain a parameter that varied across households, then optimal driving, VMT\*, would be identical for two similar households, regardless of their observed VMT. Thus, two households that have the same vehicles and household characteristics would structurally not be allowed to differ in their optimal driving. Accordingly, we non-parametrically recover  $\lambda_i$  for each household in the sample, instead of parametrically defining  $\lambda_i$  as an interaction of household characteristics, and so a singular value is recovered for each household. In order to do this, we assume that  $\lambda_i$  is a parameter that causes optimal driving to equal observed driving. In this way,  $\lambda_i$  solves the following equation:

<sup>5</sup> In order to reduce the burden on the non-linear estimation technique, we stylized the model to include only those variables that were found to be statistically significant in previous estimation routines. The results from the other estimation routines were qualitatively similar to these and are not shown here.

$$\frac{1}{J_i} \sum_{k=1}^{J_i} VMT_{ik}^* - \frac{1}{J_i} \sum_{k=1}^{J_i} \overline{VMT}_{ik} = 0 \quad (3)$$

where  $J_i$  is the number of vehicles household  $i$  owns,  $VMT_{ij}^*$  is the optimal driving as defined in equation (2), and  $\overline{VMT}_{ij}$  is the observed VMT for each vehicle  $j$  in household  $i$ 's garage (as observed in the data). Equation (3) implies that the average VMT of all vehicles in a household's garage is changed due to  $\lambda_i$ , one unique value per household, thus capturing unobserved preferences for driving that cannot be explained by mere household or vehicle characteristics.<sup>6</sup> The value of the parameter also depends on a subset of the other parameters in the model:  $\alpha$  and  $\beta$ .  $\lambda_i$  has one unique value, thus validating this quasi contraction mapping between the heterogeneity and non-heterogeneity parameters. The unique value of  $\lambda_i$  is the following:

$$\lambda_i^* = \frac{\ln\left(\sum_{k=1}^{J_i} \overline{VMT}_{ik}\right) - \sum_{k=1}^{J_i} (\alpha_{ik} + \beta_{ik} P_{ik}^u)}{\left(y_i - \sum_{k=1}^{J_i} P_{ik}^d\right)} \quad (4)$$

As  $\lambda_i$  is a function of the characteristics and other parameters, it can be thought of as a correlated random effect, which does not impede the identification of the parameters  $\alpha$  and  $\beta$ . This idiosyncratic parameter affects households' decisions on how much to drive and is a function of household and vehicle characteristics,  $\alpha$  and  $\beta$ , and observed driving.

### 3. Estimation

In our model, we allow households to optimize over a bundle of vehicles in order to explicitly model the substitution patterns among vehicles, and thus identify how relative operating costs affect VMT. However, since we have defined vehicles as a model-year-nameplate combination, this creates an extremely high dimensional choice set. The number of bundle choices available to a household are practically infinite (if allowed to choose more than one vehicle), which can hinder the ability of many discrete choice methods such as probit or logit

---

<sup>6</sup> Since  $\lambda$  shifts the average driving up or down for the household, it does not vary by the vehicles in the household's garage. However, the NHTS does not provide extra information that would help us identify a separate shifter parameter for each vehicle.

to predict decisions (as infinite choices can lead to zero probabilities and impractical estimation). Thus it is important to utilize a technique that allows for very high choice set dimensions.

The estimation technique we use is derived from Manski's (1975) semiparametric maximum score estimation strategy. This approach is based on revealed preference, which states that individuals make decisions in order to maximize their utility. Maximum score finds the parameters that maximize each individual's utility from his or her chosen outcome, given that the utility would be lower under the non-chosen outcomes. Because this technique allows for a very large choice set—as opposed to more common techniques such as logit or probit—it has been used in a wide variety of applications with very large choice sets, such as Federal Communications Commission auctions and firm location decisions.<sup>7</sup> Spiller (2011) adapts the maximum score technique by imposing distributional assumptions on the error term and including a continuous element in the decision. This adaptation of the semi-parametric maximum score approach results in a parametric maximum likelihood estimation technique, which is described in more detail below. We follow her adapted maximum score technique in order to solve this high dimensional problem.

Our strategy assumes that households have chosen their bundle of vehicles and VMT for each of their vehicles in order to maximize their utility. Furthermore, they do so given the price of gasoline and the price of their vehicle. The idea is that a household wakes up, views the price of gasoline and the prices of the vehicles, purchases these vehicles in the used vehicle market (or decides to keep its current bundle of vehicles), and decides how much VMT to allocate to each. This approach implies that if the household had made a different decision, its utility would have been lower. We can thus set up the following inequality to express this assumption:

$$V_{iB_i^*} \geq V_{iB_i} \quad \forall B_i \neq B_i^* \quad (5)$$

where  $B_i^*$  is the bundle chosen by household  $i$ . This method is optimal because it only requires the researcher to compare the chosen outcome with one other non-chosen outcome, instead of having to compare the chosen outcome with *every* non-chosen outcome as in a typical logit/probit model. The researcher thus finds the parameters that maximize this inequality for all households.

---

<sup>7</sup> Bajari and Fox (2005), Fox (2007), and Ellickson, Houghton, and Timmins (2008) utilize maximum score to solve high dimensionality problems in industrial organization settings.

However, our inclusion of vehicle-state fixed effects in the indirect utility function makes this problem even more complicated, as it would be very difficult to estimate one vehicle-state fixed effect for every combination of vehicle and state, as this would imply an explosion of the parameter space. Thus, to focus in the second stage on estimating the parameters of interest— $\alpha$ ,  $\lambda$ ,  $\beta$ , and  $\tau$ —it is important to eliminate the fixed effects in a first stage. With this in mind, we set up a bilateral inequality across households instead of comparing only within households. As an example, consider two households, A and B, who have vehicles 1 and 2, respectively. Equation (5) implies that for household A, owning vehicle 1 is better than owning vehicle 2. Similarly, for household B, owning vehicle 2 is better than owning vehicle 1. Thus we can rewrite (5) in the following way:

$$\begin{aligned} V_{A1} &\geq V_{A2} \\ V_{B2} &\geq V_{B1} \end{aligned}$$

We can expand the indirect utility function in order to separate out the vehicle-state fixed effects:

$$\begin{aligned} \bar{V}_{A1} + \theta_{S1} &\geq \bar{V}_{A2} + \theta_{S2} \\ \bar{V}_{B2} + \theta_{S2} &\geq \bar{V}_{B1} + \theta_{S1} \end{aligned}$$

where  $\bar{V}$  is the variable portion of the indirect utility function that does not include the vehicle-state fixed effect  $\theta_{Sj}$ . We can add across these two inequalities, yielding a third inequality:

$$\bar{V}_{A1} + \theta_{S_{A1}} + \bar{V}_{B2} + \theta_{S_{B2}} \geq \bar{V}_{A2} + \theta_{S_{A2}} + \bar{V}_{B1} + \theta_{S_{B1}}$$

By doing this sort of bilateral comparison for two households from the same state (where  $S_A = S_B$ ), the vehicle-state fixed effects  $\theta_{S_{A1}}$  and  $\theta_{S_{B2}}$  drop out of both sides of the equation, thereby eliminating the fixed effects and leaving one inequality that only depends on  $\alpha$ ,  $\lambda$ ,  $\beta$ , and  $\tau$ , the parameters of interest:

$$\bar{V}_{A1} + \bar{V}_{B2} \geq \bar{V}_{A2} + \bar{V}_{B1} \quad (6)$$

Since the fixed effects are state specific, this only works if the two households are located in the same state, otherwise the fixed effects would not drop out of both sides. Thus, we implement a “swap” that evaluates the utility function of two households in the same state under an alternative vehicle (or vehicle bundle). Specifically, this alternative vehicle is “swapped” between the two households.

While maximum score would find the parameters that maximize the number of times that this equation holds for each household, we take a maximum likelihood approach and utilize the Normality assumption on the error term in order to create a better behaved likelihood function.<sup>8</sup> We can rewrite Equation (6) again by separating out the error terms associated with the indirect utility function:

$$\tilde{V}_{A1} + \varepsilon_{A1} + \tilde{V}_{B2} + \varepsilon_{B2} \geq \tilde{V}_{A2} + \varepsilon_{A2} + \tilde{V}_{B1} + \varepsilon_{B1}$$

where  $\tilde{V}$  is the deterministic portion of the indirect utility function that does not include the vehicle-state fixed effect or the error term. As the summation of normal error terms is also equal to a normal error, we can group the error terms into one solitary error:

$$\tilde{\varepsilon} \leq \tilde{V}_{A1} - \tilde{V}_{A2} + \tilde{V}_{B2} - \tilde{V}_{B1}$$

where  $\tilde{\varepsilon}$  is the summation of all four  $\varepsilon$  terms. Given the normality of  $\tilde{\varepsilon}$ , we can find the probability that any random swap will make both households worse off:<sup>9</sup>

$$P(\tilde{\varepsilon} \leq \tilde{V}_{A1} - \tilde{V}_{A2} + \tilde{V}_{B2} - \tilde{V}_{B1}) = \Phi\left(\frac{\tilde{V}_{A1} - \tilde{V}_{A2} + \tilde{V}_{B2} - \tilde{V}_{B1}}{2\sigma}\right) \quad (7)$$

We estimate Equation (7) via maximum likelihood in Matlab, using a gradient-based technique. With this approach, we find the parameters that maximize the probability that every household made an optimal decision with respect to the vehicle purchase and VMT decisions, and the assumption that they would have been worse off had they made a different choice.

#### 4. Data

The data we use for estimation come from various sources. The main dataset is the 2009 National Household Travel Survey (NHTS). This survey provides a national dataset of

---

<sup>8</sup> Parameterizing the error term and estimating through maximum likelihood also allows us to find point estimates on the parameters instead of ranges of values, which would be the result of maximum score. This allows for easier policy analysis.

<sup>9</sup> Since each error term  $\varepsilon_{ij} \sim N(0, \sigma^2)$ , the summation over the four error terms associated with equation (7) will be distributed normally with a variance equal to  $4\sigma^2$ .

households, their demographic characteristics, location, vehicles owned, and driving patterns.<sup>10</sup> We supplement the publicly available data by obtaining confidential data from the NHTS that locates households at the census tract level,<sup>11</sup> providing us with much finer detailed geographic location information. The 2009 NHTS dataset is comprised of 309,163 vehicle observations, owned by 143,084 households across the United States. Our data-cleaning process to remove observations with missing data reduced our final sample to 77,637 households.<sup>12</sup> Summary statistics of the full NHTS dataset, our estimation sample, and a comparison with 2010 Census data can be found in Table 3. While the median income in the final sample is somewhat above the national median income, it appears more like the nation than the overall NHTS 2009 dataset. Besides income, though, the full NHTS and the final estimation sample appear to be relatively similar, providing some evidence that there was not any systematic reason for why some observations had missing data.

**Table 3. Summary Statistics NHTS**

	NHTS 2009 (in means)	Final sample (in means)	2010 Census (in means)
Number of observations	309,163	139,115	308,745,538
Number of households	143,084	77,637	116,716,292
Household income (median)	\$62,500	\$52,500	\$ 49,445
Household size	2.62	2.48	2.58
Percent urban	77%	71%	80%
Percent white	87%	88%	75%
Household vehicle count	2.71	1.79	-
Household fuel efficiency	21.13	21.68	-
Vehicle age	9.24 years	7.62 years	-
Vehicle miles traveled	9,729	10,248	-

<sup>10</sup> See Table 2 for a discussion of which variables are included in the model.

<sup>11</sup> Non-confidential data gives location at the metropolitan statistical area (MSA) level, but only if the MSA is large enough (1 million or more).

<sup>12</sup> One of the main reasons that the overall number of households in the sample was reduced was if we were unable to match one of the household's reported vehicles to our vehicle dataset: if a household had missing information on one of its vehicles, we had to remove the entire household from the sample.

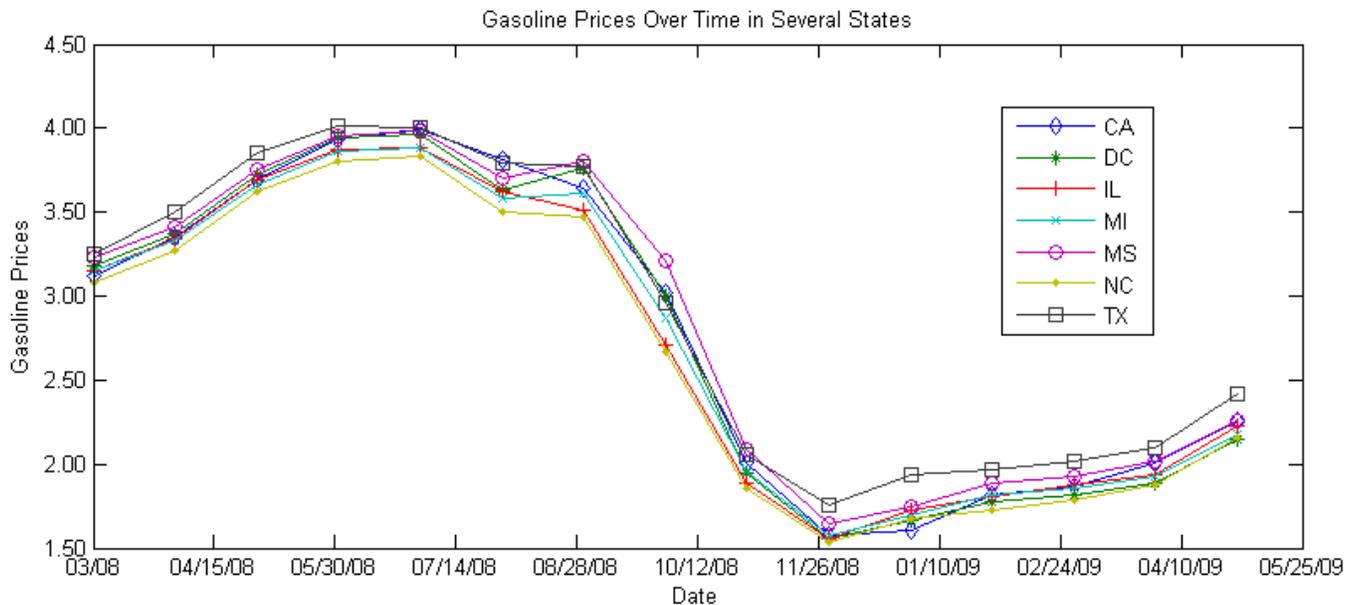
We collected gasoline prices at the state-month level, and assigned these prices according to the date of the household's survey, which is reported in NHTS. These prices were gathered from the U.S. Energy Information Administration (EIA 2011) and supplemented with federal and state tax information from EIA, as well as with additional state tax information from individual states (US FHWA 2009a and 2009b). The state tax information includes variable percentage gasoline taxes, flat taxes, and sales taxes that affect the final purchase price. Together with the federal taxes, we are thus able to calculate a final estimated monthly state gasoline price per gallon. While we recognize that there is within-state variation in prices, there is no reason to believe it is biased, and Hastings (2010) shows that localized data may be biased because of the way in which it is collected. A sample range of different gasoline prices across states is summarized in Table 4.

**Table 4. Gasoline Prices Across States and Time**

<b>State</b>	<b>Gasoline price</b>	<b>Date</b>
<b>Vermont</b>	\$1.45	January, 2009
<b>Arizona</b>	\$1.62	December, 2008
<b>Tennessee</b>	\$2.20	November, 2008
<b>Delaware</b>	\$2.30	May, 2009
<b>Illinois</b>	\$3.15	March, 2008
<b>Iowa</b>	\$3.68	September, 2008
<b>California</b>	\$3.99	July, 2008
<b>Arkansas</b>	\$4.55	June, 2008

Figure 1 shows the change in gasoline prices over time during March 2008 to May 2009 for seven different states—California, Washington DC, Illinois, Michigan, Mississippi, North Carolina, and Texas. The gas prices track closely across states, as they all peak in the summer of 2008 and dip back down during 2009. Thus, there is substantial variation in gasoline prices across states and over time, allowing us to identify the elasticity of demand for gasoline of different households depending on the location and date of the survey.

Figure 1. Gasoline Prices over Time and States



We merged the vehicles in the NHTS dataset with vehicle characteristics, which were hand entered from Ward's Automotive Yearbooks. We also obtained used vehicle prices based on October 2009 average used vehicle transaction data from the National Automotive Dealers Association (NADA). However, since the NHTS 2009 survey took place over a year-long window (March 2008–May 2009), we use these vehicle prices to predict the used vehicle prices during the month the survey was conducted in a household given that month's average state gasoline price. This implies that each household would face a different used vehicle price depending on geographic location and date of the survey. The hedonic price gradient used to predict these prices is shown in equation (8).

$$P = X\beta + \varepsilon \quad (8)$$

where  $X$  is a vector of characteristics including operating costs, a high mileage dummy, vehicle age\*MPG, vintage dummies, MSRP, length\*width, weight, horsepower, and a variety of manufacturer dummies that enter the equation both individually and interacted with premium and large vehicles. This captures the fact that used vehicle prices change with the age of the vehicle as well as the gasoline prices, thus adjusting for the fact that October 2009 prices may not adequately reflect the used price of the same vehicle in March 2008, given the wide variation in

gasoline prices during 2008–2009 and the difference in vehicle age. Equation (8) is estimated through ordinary least squares.

Our data also include census tract-level information such as distance to the population-weighted center of the nearest MSA, which is calculated using Arc-GIS, census tract population density, and county-level information on median rents in order to control for cost of living.<sup>13</sup> Median rents are based on Fair Market Rents from the U.S. Department of Housing and Urban Development.

## 5. Results

The “swapping” estimation technique allows us to identify the preference parameters that tell us how different households value different vehicle characteristics. Once these preference parameters are estimated, we calculate optimal VMT per household under original gasoline prices and a 1 percent increase in gasoline prices. We use the difference in VMT under these two different gasoline price schedules to find the elasticity of demand for gasoline. The elasticity of each household is calculated as:

$$\varepsilon_i = \frac{\sum_{j_i} ((VMT_{j_0} / mpg_j) - (VMT_{j_1} / mpg_j))}{price_0 - 1.01 * price_0} * \frac{price_0}{\sum_{j_i} (VMT_{j_0} / mpg_j)}$$

where  $price_0$  is the initial gasoline price faced by household  $i$ ,  $VMT_{j_0}$  is the initial optimal VMT of vehicle  $j$  in household  $i$ 's bundle of vehicles (given the initial gasoline price), and  $mpg_j$  is the miles per gallon of vehicle  $j$  in household  $i$ 's bundle of vehicles. Thus, the numerator on the left is the change in overall gallons demanded by household  $i$ , and the denominator on the right is the original gallons demanded by household  $i$ . This calculation produces one elasticity estimate per household, allowing for a rich analysis of how elasticities vary across different locations and demographics. This is a short-term elasticity estimate, as we do not allow households to re-optimize over their bundle of vehicles given an increase in gasoline prices.<sup>14</sup>

<sup>13</sup> Census tract-level estimates of housing values and rents are only available for all areas using the 2005–2009 five-year estimates from the American Community Survey.

<sup>14</sup> The method implemented here does not allow for a straightforward estimation of the long-run elasticity and would need to be adapted significantly. This is outside the scope of this paper and thus remains for future work.

We average over all elasticities in order to find a mean elasticity across the entire sample of households.<sup>15</sup> The results from the estimation and the resulting average elasticity estimate are presented in Table 5.

The signs of the parameter estimates are fairly intuitive with high levels of significance. For example, households living in densely populated areas prefer shorter vehicles, presumably so that it is easier to park in these dense areas. Other parameters are also intuitive; households that have farther to travel to work or access urban centers prefer more efficient vehicles, and richer households prefer newer and faster vehicles. The coefficient on vehicle size\*household size is negative, which may seem counterintuitive in that we would expect to see larger households prefer larger vehicles. However, upon further examination of the data, we see that the average household size is just slightly higher than two. Thus, a marginal change in household size presumably would be because of either one extra child or an extra driver. If it is an extra driver, then there is no need to have a larger vehicle.

**Table 5. Estimation Results and Elasticity**

	Parameter	Coefficient (standard error)
$\alpha$	Household size	0.0415***(0.003)
	Income/median rent	0.0106***(0.0001)
	Average commute time	0.6617***(0.002)
	# Drivers/ # vehicles	0.1296***(0.004)
	Rural location dummy	0.0931***(0.003)
$\beta$	Vehicle size*household size	-0.4739*** (0.003)
	Acceleration*relative income	0.0205*** (0.001)
	Truck*distance to MSA	-0.0133***(0.001)
	MPG*commute distance	0.0079***(0.000)
$\tau$	Vehicle age*income	-0.0146***(0.002)
	Vehicle length*population density	-1.5903***(0.017)
$\sigma^2$	(Exp) variance of error term	2.2476*** (0.005)
	Implied average elasticity	-0.6679 (0.6063)
(Statistical significance: *** 99%, ** 95%)		

<sup>15</sup> We remove 1.5 percent of outliers from the top and the bottom in order to not skew the results to extreme data points.

As we cannot directly test the validity of  $\lambda$  through parameter values and significance, we do so indirectly by measuring the correlation between optimal and observed driving. Our correlation parameters are large: the relationship between average optimal and observed VMT within each household is 0.718, and the relationship between the optimal and observed driving for each vehicle separately is 0.675. This high level of correlation provides a form of verification to the model in that our constructed optimal driving reflects the data. Thus,  $\lambda$  allows us to capture a large portion of the factors that affect driving and to allow for unobserved heterogeneity across households.

The elasticity estimate of -0.67 is also reasonable and is near the average elasticity estimated in the literature (-0.5). As we have one elasticity estimate per household, we analyze how the elasticities vary across different household characteristics and present the results in Table 6. We find that households with higher incomes are more elastic. Furthermore, households with more vehicles are more elastic, proving that substituting between vehicles in the garage allows the household to react more readily to gasoline prices. Households facing higher gasoline prices and higher VMT are also more elastic. These types of households have a higher share of consumption dedicated to driving, and thus any increase in gasoline prices will be felt more acutely by these individuals.

The interesting story is when it comes to distances to MSAs and rural locations. Table 6 shows that living far from an urban center and being in a rural household increases a household's elasticity. However, these are the households that drive the most. Predicted optimal driving and observed VMT for households in rural areas as well as those that live farther from urban centers are much higher than their urban city-dwelling counterparts. Thus, it is not surprising that these households are much more sensitive to gasoline prices. However, one might argue that those who live in urban areas have better access to public transportation and thus should be more elastic than rural households. While this is a valid argument, it is possible that those who live in remote, rural areas drive so much more, on average, that they can cut out one trip and see a large change in their driving. For example, if a household usually makes three separate trips into the urban center per week, a change in gasoline prices may cause this household to group these trips into one or two weekly visits. This would greatly minimize VMT and thus they would have a larger elasticity than those who live in urban centers. Presumably, people living in urban centers already optimize their driving needs or do not drive as much and thus make relatively small changes in their driving habits (i.e. walking to the store instead of driving), which will decrease overall VMT less than for rural households. If we were able to control for quality of public

transit, this stark difference between rural and urban household elasticities might decrease. However, because there is currently no good data on public transit availability with wide enough coverage nationally, we are not able to distinguish the household's ability to substitute to public transit as gasoline prices increase.<sup>16</sup>

**Table 6. Elasticity by Household Characteristics**

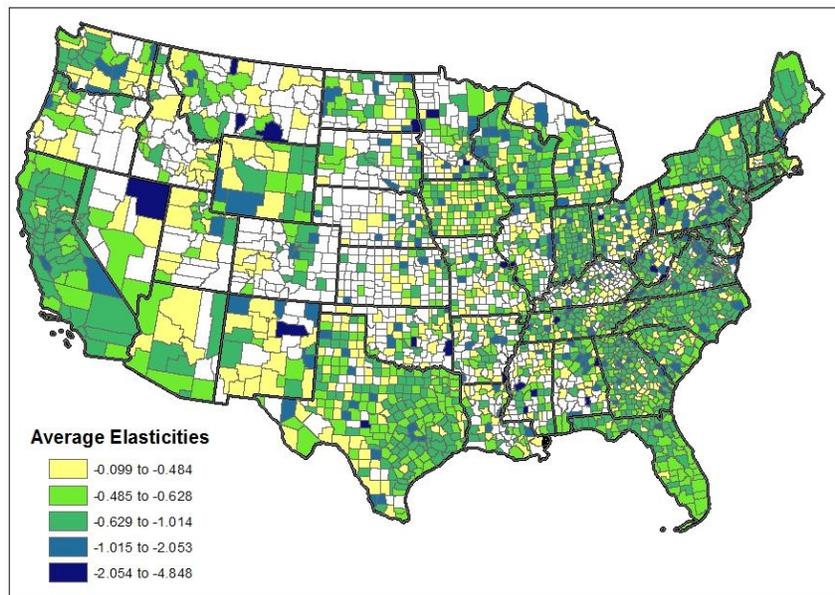
(Averages)	>-0.25	-0.25:-0.35	-0.35:-0.45	-0.45:-0.6	-0.6:-0.75	-0.75:-1	-1:-1.5	>-1.5
Distance to MSA (in meters)	37,709	40,277	40,930	41,382	41,420	40,490	40,939	43,060
Income/median rent	0.58	0.58	0.64	0.63	0.67	0.78	0.82	0.88
Income	4.92	4.82	5.39	5.24	5.57	6.57	6.95	7.63
# Vehicles	1.42	1.47	1.66	1.66	1.84	2.10	2.32	2.51
Commute time	17.22	17.27	18.34	20.23	20.72	22.35	26.63	36.60
Gasoline price	2.16	2.32	2.63	2.99	3.31	3.31	3.35	3.34
MPG	22.96	22.05	21.54	22.00	21.10	21.07	21.40	22.22
VMT	14,452	13,299	15,498	15,528	17,471	21,735	25,311	32,509
% Rural	0.21	0.24	0.24	0.25	0.28	0.29	0.33	0.38
# Obs.	10,805	12,646	10,217	11,887	9,921	7,641	6,639	5,389

To further explore the heterogeneity across both households and regions, in Tables 7–9 we repeat the analysis of elasticities by household characteristics but separate the sample into three different geographic regions (based on the Census Division): Pacific, Midwest, and Southeast. We see that there is significant heterogeneity in household characteristics and how they affect elasticity across the regions. For example, the Pacific Census Division has far fewer rural households, and the elasticities are well distributed across these rural households. Thus, it does not appear that the negative relationship between elasticity and rural location is necessarily the same in the Pacific versus the Midwestern and Southeastern regions. Furthermore, in the Pacific region, elasticity is relatively similar across different distances to the nearest MSA, while in the other two regions, living farther from an MSA increases household elasticity.

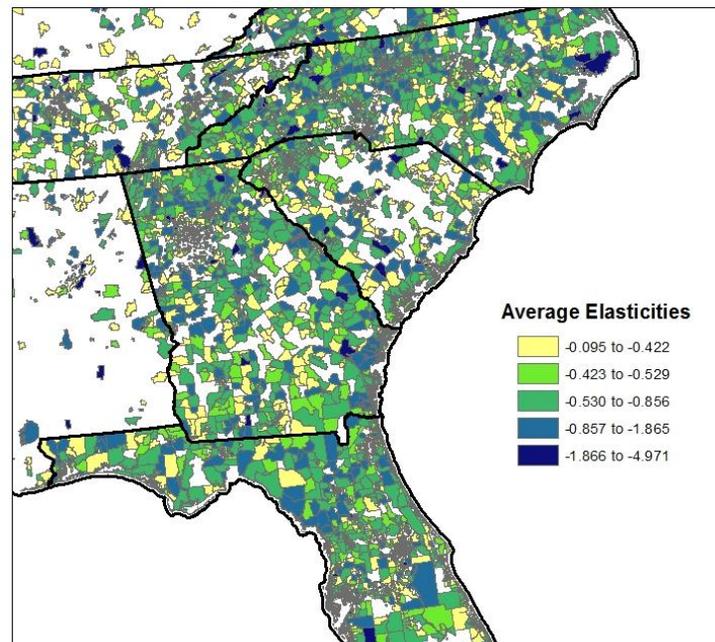
<sup>16</sup> Spiller et.al. (2012) construct a measure of public transit substitutability (or quality) but it is only valid for individuals living in urban areas; thus, their method would not work for rural households or be applicable in this analysis.

This heterogeneity across regions can also be seen in Figure 2, which presents a national map of elasticities (household elasticities are aggregated at the county level), and Figure 3, which shows a closer look at the southeastern region of the United States (household elasticities here are aggregated at the census tract level).<sup>17</sup> Even in southern Georgia, where there is a very large percentage of rural households, there is a tremendous amount of heterogeneity in the elasticities.

**Figure 2. Average Elasticities by County**



<sup>17</sup> Figure 3 shows more variation across census tracts than is reflected in Figure 2, with some adjacent census tracts having extremely different average elasticities, yet this could be due to possible small numbers of observations at the census tract level.

**Figure 3. Southeastern US Census Tract Level Elasticities**

The results demonstrate that it is important to take into account regional differences when implementing public policies that affect gasoline prices and driving. One way to do this would be to apply state level policies instead of national, one-size-fits-all tax policies. Therefore, we demonstrate that increasing gasoline prices, such as through increasing gasoline taxes, are likely to result in diverse welfare impacts based on geographic location. Thus, in the next section, we analyze the impact on households given a national 10 cent per gallon gasoline tax in order to better understand its welfare impacts. Furthermore, we propose a revenue recycling method that can help to counteract the negative welfare impacts of the tax on critical communities (such as people with low incomes).

**Table 7. Elasticity by Household Characteristics, Pacific Census Division**

(Averages)	>-0.25	-0.25:-0.35	-0.35:-0.45	-0.45:-0.6	-0.6:-0.75	-0.75:-1	-1:-1.5	>-1.5
Distance to MSA (in meters)	45,653	46,169	45,643	45,973	45,153	44,745	44,996	44,906
Income/median rent	0.42	0.42	0.46	0.48	0.47	0.57	0.61	0.66
Income	5.20	5.31	5.70	5.98	5.89	7.14	7.72	8.29
# Vehicles	1.45	1.44	1.63	1.68	1.75	2.00	2.21	2.45
Commute time	18.31	19.97	19.69	21.80	22.08	23.31	27.75	36.56
Gasoline price	2.41	2.52	2.84	3.22	3.57	3.67	3.64	3.65
MPG	24.36	23.41	22.38	23.03	22.25	21.70	22.06	23.24
VMT	14,044	12,654	14,252	15,051	15,619	20,067	23,156	29,906
% Rural	0.07	0.09	0.09	0.08	0.09	0.10	0.10	0.13
# Obs.	1,389	1,404	1,487	1,552	1,423	1,310	964	933

**Table 8. Elasticity by Household Characteristics, Midwest<sup>18</sup>**

(Averages)	>-0.25	-0.25:-0.35	-0.35:-0.45	-0.45:-0.6	-0.6:-0.75	-0.75:-1	-1:-1.5	>-1.5
Distance to MSA (in meters)	30,438	32,949	35,727	34,577	34,527	35,659	34,294	37,491
Income/median rent	0.66	0.61	0.69	0.68	0.71	0.90	0.91	1.05
Income	4.77	4.38	4.93	4.89	5.06	6.38	6.44	7.42
# Vehicles	1.45	1.48	1.67	1.63	1.79	2.13	2.33	2.57
Commute time	16.58	15.23	16.02	19.02	19.06	20.72	24.53	33.13
Gasoline price	2.22	2.24	2.67	3.03	3.37	3.38	3.38	3.41
MPG	22.11	21.66	21.15	22.15	20.86	20.80	21.33	21.69
VMT	14,316	12,881	15,222	14,978	16,448	21,802	27,259	32,305
% Rural	0.20	0.23	0.26	0.23	0.29	0.31	0.37	0.43
# Obs.	462	626	480	566	549	386	341	286

<sup>18</sup> This includes the East North Central Census Division.

**Table 9. Elasticity by Household Characteristics, Southeast<sup>19</sup>**

(Averages)	>-0.25	-0.25:-0.35	-0.35:-0.45	-0.45:-0.6	-0.6:-0.75	-0.75:-1	-1:-1.5	>-1.5
Distance to MSA (in meters)	32,435	34,963	34,113	34,631	36,224	35,801	36,573	40,156
Income/Median rent	0.58	0.60	0.67	0.64	0.70	0.81	0.86	0.94
Income	4.66	4.71	5.31	5.00	5.36	6.24	6.63	7.32
# Vehicles	1.39	1.49	1.70	1.68	1.88	2.16	2.40	2.58
Commute time	16.36	16.26	17.20	19.48	19.99	21.47	25.36	36.33
Gasoline price	2.05	2.24	2.55	2.92	3.25	3.22	3.30	3.25
MPG	23.06	21.88	21.38	21.77	20.81	20.94	21.19	21.81
VMT	13,700	13,100	15,800	15,300	17,200	22,000	25,400	33,500
% Rural	0.23	0.26	0.27	0.30	0.33	0.35	0.40	0.46
# Obs.	4,236	5,123	3,757	4,578	3,735	2,778	2,466	1,920

## 6. Welfare Analysis & Policy

In order to understand how the welfare impact of a national gasoline tax varies across households, we calculate the equivalent variation (EV) and the compensating variation (CV) as a percentage of household income given a 10 cent per gallon tax. We find that the average amount of money needed to reimburse households so that they stay at the same utility as prior to the increase (CV) is 0.8 percent of total household income. The EV (the amount the household is willing to spend to avoid the increase) relative to income is similar, but slightly larger: 0.9 percent. This is, on average, \$484 for EV and \$384 for CV, which is significantly more than the amount of money an average household in the sample would have spent on a 10 cent per gallon gasoline tax (\$86).<sup>20</sup> We also find that both relative EV and CV are higher for rural households. The relative EV of a 10 cent gasoline tax for rural households is 1.2 percent versus 0.89 percent for urban households; the relative CV for rural households is 1 percent versus 0.75 percent for urban households. This demonstrates that rural households are hurt almost 30 percent more by gasoline taxes than urban households. Even though rural households, on average, have higher elasticities, they drive considerably more miles than do urban households and thus have higher welfare effects from higher prices.

<sup>19</sup> Southeast Census Divisions include East South Central and South Atlantic.

<sup>20</sup> We run the analysis with different levels of price increases and find quantitatively similar results, with respectively larger responses at higher gasoline price increases.

Table 10 demonstrates how the relative size of the CV measure (as a percentage of income) changes with household characteristics.<sup>21</sup> The CV percentages are increasing in distance to MSA, number of vehicles, VMT, commute time, and percent rural. These households are affected more than others by higher gasoline prices, given that they demand more driving. Furthermore, CV decreases in miles per gallon, implying that those with more efficient vehicles are better able to respond to higher gasoline prices and thus do not suffer as much from negative welfare effects as do others.

The effect of income appears to be U-shaped in that poorer households are located at the high and low ends of the CV percentage spectrum; however, this may simply be due to the heterogeneity of poor households—some may drive very little to save money, while others may commute long distances for low-paying jobs.

Figure 4 shows the variation across different regions of the United States of relative CV (aggregated to the county level) given a 10 cent per gallon tax. There is tremendous heterogeneity across and within states, further demonstrating how widely the welfare impacts vary.

---

<sup>21</sup> We conduct the same analysis with EV and since the results are qualitatively the same we do not produce them here.

Figure 4. Average Compensating Variation as a Percentage of Income

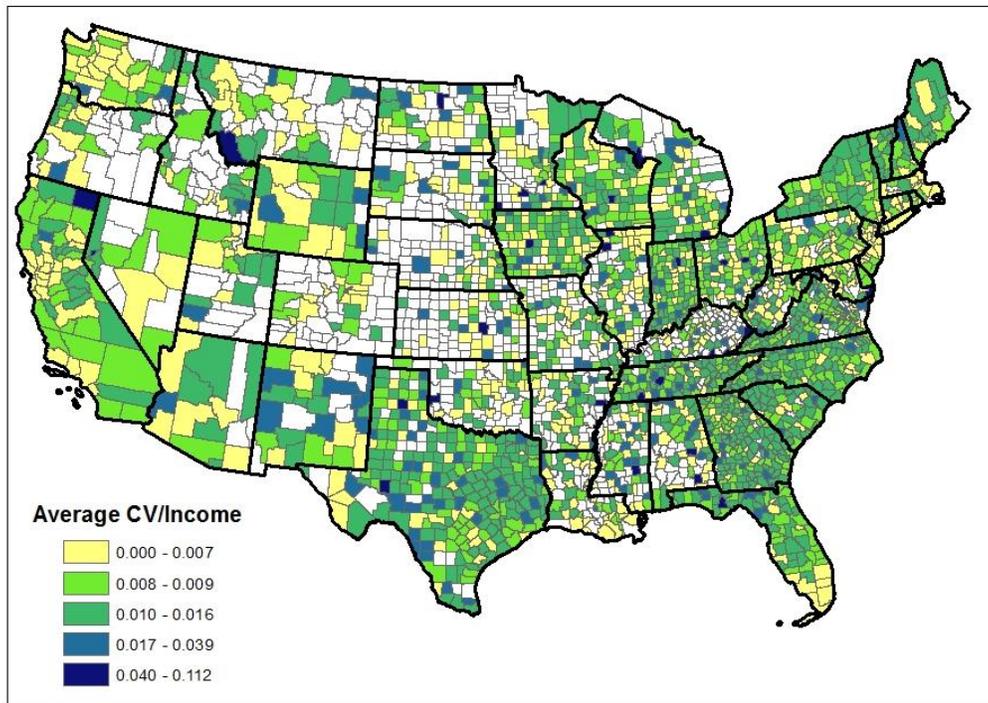


Table 10. Compensating Variation as Percentage of Income

(Averages)	<.25%	.25:.4%	.4:.55%	.55:.7%	.7:.9%	.9:1.3%	1.3:2.3%	>2.3%
Distance to MSA (in meters)	39,087	39,118	38,366	39,431	40,131	42,679	45,379	46,965
Income/median rent	0.61	0.71	0.76	0.78	0.74	0.64	0.50	0.30
Income	54,500	63,300	66,100	65,900	60,500	50,700	37,600	22,200
# Vehicles	1.29	1.58	1.79	1.94	2.00	2.01	1.92	1.73
Commute time (mins)	21.11	22.84	23.67	23.87	23.51	22.61	19.94	16.26
Gasoline price	2.86	2.84	2.83	2.81	2.82	2.85	2.89	2.94
MPG	23.47	22.92	22.19	21.63	21.22	20.78	20.56	20.34
VMT	3,830	9,643	14,614	19,151	22,592	25,673	28,908	36,579
% Rural	0.16	0.18	0.20	0.25	0.30	0.36	0.41	0.42
# Houses	10,850	10,382	11,855	10,669	10,494	10,577	7,135	2,993

We now demonstrate that it is possible for the government to mitigate the disparate welfare impacts of a gasoline tax by recycling the revenue from the tax back to households.

Given a 10 cent gasoline tax, we propose three different revenue recycling mechanisms, with each reimbursing a certain percentage of overall welfare losses depending on the way in which the refund is returned.

The first option (Policy 1) is to refund the revenue in a lump sum fashion, and in equal parts to each household, regardless of welfare impacts. In this way, the government would send a check for \$86.44 to each household. This check, on average, would cover approximately 67 percent of the CV for each household, though how much of the CV is refunded varies widely, with approximately 14 percent of households receiving a refund of more than 100 percent of their CV.<sup>22</sup> Given that this flat refund provides some households with more income than would be necessary to compensate them for welfare effects from the higher gasoline taxes, this would be an inefficient refund.

In order to make the refund more efficient, the refund could be based on certain household characteristics. However, Table 10 shows that certain households are more impacted by gasoline prices because of decisions that make them more likely to use more gasoline, such as choosing less-efficient vehicles or commuting long distances. Thus it would not be socially beneficial to compensate these households based on these characteristics. Therefore, we propose that the government could choose to refund money based solely on characteristics such as income, household size, and location (Policy 2).

We first calculate the optimal refund as the share of each household's CV multiplied by the total revenue received from the federal gasoline tax (10cents per gallon).

$$Share_i = \frac{CV_i}{\sum_{i=1}^N CV_i}$$

$$refund_i = revenue * Share_i$$

where  $i=1 \dots N$  is the number of households in our sample.

The revenue to be diverted for refunds is 10 cents multiplied by the total number of gallons of gasoline demanded under a 10-cent-higher gasoline price. Once the refund is calculated, we regress the calculated refund on household location (including distance to MSA, census division dummies, and a rural dummy), income, household size, gasoline price, and

---

<sup>22</sup> This average of 67 percent is inflated by individuals with very low CV values. For those receiving refunds less than their CV, the refund covers less than 30 percent.

second orders of the continuous parameters.

$$refund = \beta_0 const + \beta_1 DIST\_MSA + \beta_2 HH\_SIZE + \beta_4 GAS\_P + \beta_5 PACIFIC + \beta_6 MOUNTAIN + \beta_7 MIDWEST + \beta_8 RURAL + \beta_9 DIST\_MSA^2 + \beta_{10} HH\_SIZE^2 + \beta_{11} GAS\_P^2 + \varepsilon$$

The results from this regression would allow the government to predict refunds based only on household characteristics and location without needing to calculate CV for each household. Similar to how tax refunds are assigned through tax returns, households could input this information in the tax return at the end of the year and receive a credit depending on their location and characteristics.

We predict refunds given the estimated coefficients. The resulting average refund for Policy 2 is \$70.03, which is slightly lower, both in levels and as a percentage of CV (34.8 percent), than Policy 1, the flat refund. However, the distribution of refunds is more efficient, in that only 4.5 percent of households have more than 100 percent of their CV returned to them. Furthermore, the amount that the return exceeds the CV for this small group of households is lower under this form of revenue recycling compared to a flat tax (on average \$18 versus \$37).

**Figure 5. Difference Between CV and Refund, Given Different Recycling Mechanisms**

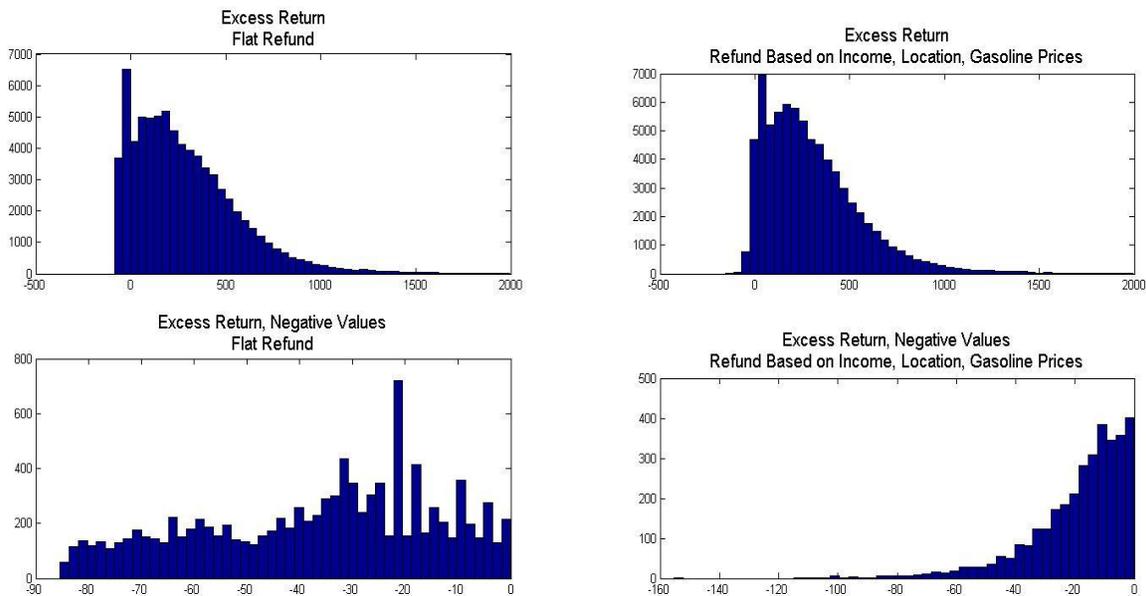


Figure 5 demonstrates the difference between each household’s CV and its refund under the two different recycling mechanisms. The graphs on the right show the excess return (CV

minus refund) when the refund is calculated using household characteristics; the left graphs assume a flat refund to all households. The bottom row shows the distribution of excess returns for households that receive a return greater than their CV. This figure demonstrates that although the average difference between the CV and refund amount is relatively similar under the two mechanisms, the flat return is less efficient and has a higher distribution of negative differences (or more people who receive a refund much higher than their welfare losses).

However, one issue with Policy 2 is that the majority of the reimbursement will go to those with higher incomes, given that these households are more elastic and thus face the highest welfare losses. In Figure 6, we plot the distribution of returns by income for the second revenue recycling mechanism and the distribution of relative CV to income.

**Figure 6. Equity Concerns of Revenue Recycling in Policy 2**

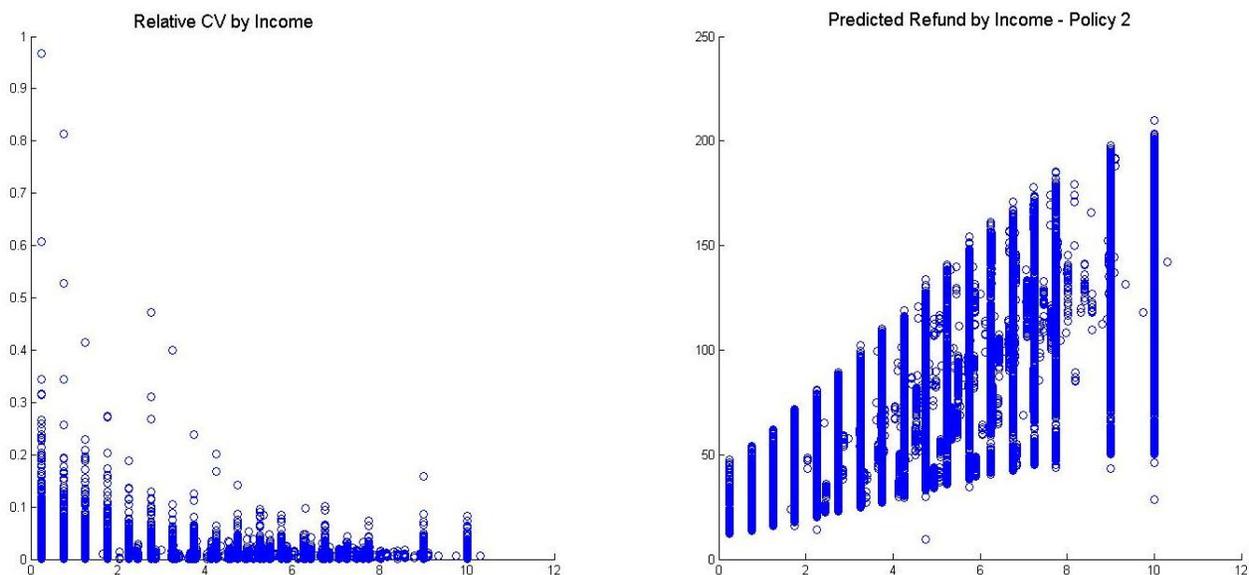


Figure 6 demonstrates that although relative CV (the welfare impacts given the household's income) is decreasing by income, implying that poor households are more impacted by the gasoline tax, the proposed refund under Policy 2 is increasing in income. To address this inequity, we thus propose a third policy of revenue recycling. Under Policy 3, we implement a similar technique for predicting refunds as in Policy 2, but instead of calculating the share as a function of CV, the share is now based on relative CV (or CV as a share of income). Thus:

$$Share_{i2} = \frac{CV_i / income_i}{\sum_{i=1}^N (CV_i / income_i)}$$

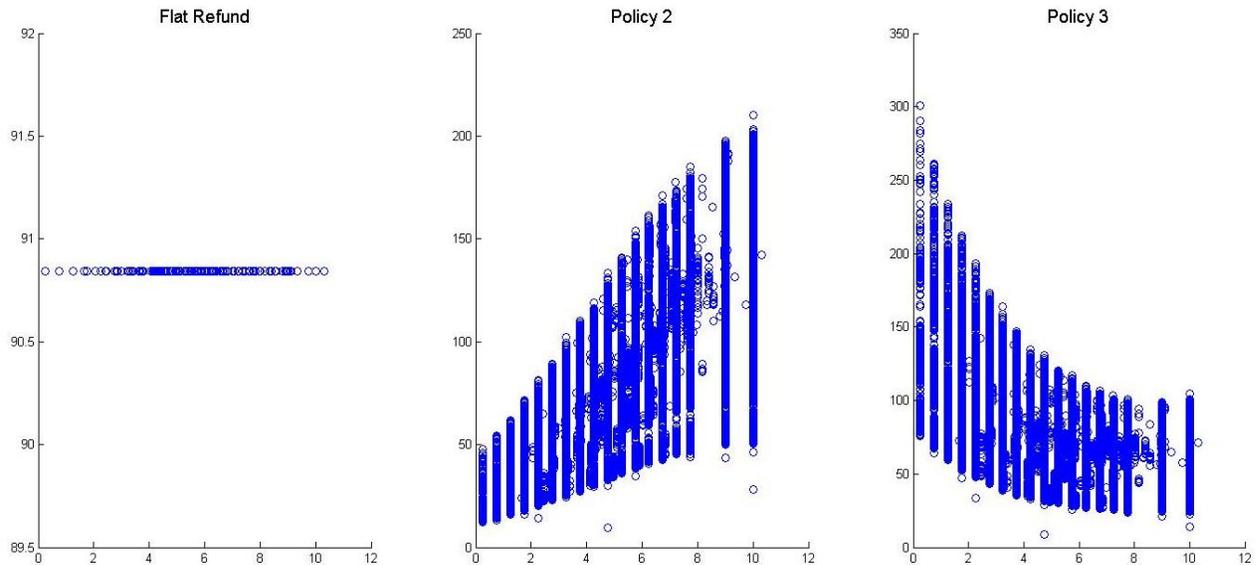
and

$$refund_{i2} = revenue * Share_{i2}$$

This allows the refund to be based on relative CV, thus accounting for the fact that lower income households face welfare effects that are a greater percentage of their income. The average refund under this policy is \$64.07, which is on average 53 percent of total CV. While this policy is not as efficient as Policy 2 (9.5 percent of households receive, on average, refunds that are \$37 above their CV), it is more efficient than Policy 1 and much more equitable. Figure 7 plots the predicted refund by income for all three policies, showing that Policy 3 favors poorer households.

Therefore, we find that by implementing Policy 3, the government can implement a gasoline tax to reduce driving and reduce gasoline consumption, while also offsetting, on average, 53 percent of the welfare losses associated with that tax and doing so in an equitable manner.

Policy 3 also retains some of the gasoline tax as revenue. While the first policy, the flat tax, results in no revenue for the government because all proceeds from the gasoline tax are recycled to households, the other two policies raise revenue because not all of the tax revenues are returned to households. Furthermore, Policy 3 results in approximately 37 percent more revenue than Policy 2. (The government keeps 2.6 cents per gallon under Policy 3, while under Policy 2, it only keeps 1.9 cents per gallon.) Thus, through Policy 3, the government is able to refund a significant amount of money to households to mitigate some of the welfare losses due to higher gasoline prices while still accruing a positive revenue stream. Furthermore, this policy will help induce households to internalize some of the externalities associated with gasoline consumption, thus reducing greenhouse gases and U.S. reliance on foreign oil.

**Figure 7. Equity Concerns Across All Three Revenue Recycling Policies**

## 7. Conclusion

Combating the negative externalities of gasoline consumption through the implementation of a gasoline tax requires understanding how consumers will respond—and not just at the national level. As we have found, there is tremendous heterogeneity in the responses among different types of households and different regions. Additionally, higher gasoline prices can impose large, negative welfare effects on households with high VMT needs, even those with high gasoline price elasticities.

We demonstrate that households' gasoline elasticity varies with multiple characteristics and demographics, including the income they earn, the number of vehicles they own, the average annual number of miles they drive, and the amount of distance they must travel to reach an urban area. Across regions we see further evidence of the heterogeneity in households' ability to respond to increases in gasoline prices. For example, in the Pacific Census region, there is tremendous heterogeneity in the elasticities of rural households, while in the Midwest, more rural areas tend to have higher elasticities.

In general, however, policymakers are less concerned with elasticities than in the actual welfare effects on households. As is the case with the household-level elasticities, gasoline taxes have heterogeneous effects on household welfare, as shown by the map in Figure 4. While rural

households appear to have more elastic gasoline demand, they also drive more miles. Thus, in our welfare analysis we find that a 10 cent increase in gasoline prices would have a 30 percent larger negative welfare effect on rural households versus their urban counterparts. And as Weber and Jensen (2004) find, rural areas are more likely to have higher levels of persistent poverty, making it even more difficult for these areas to cope with welfare losses.

We also find that the number of vehicles increases household elasticity. This holds with the theory that households with more than one vehicle are better able to respond to higher gasoline prices by substituting toward more fuel efficient vehicles within their bundle. Thus, the number of vehicles helps households cope with rising gasoline prices. However, this does not necessarily translate into higher welfare for multivehicle households: these households still may have to drive more miles than will a household with fewer vehicles if they live farther from urban amenities or have longer commute times.

We examine several policy scenarios that the government could implement to mitigate some of these negative welfare impacts. We find that by recycling the tax revenue based on the characteristics associated with the relative welfare effects from higher gasoline prices, the government can help those most affected by rising gasoline prices and could reduce the welfare impact on average by 53 percent. Additionally, unlike a flat revenue recycling policy, this program would still generate positive government revenues from the tax.

It is possible that our current analysis does not account fully for the heterogeneity of how household characteristics and location parameters affect household utility. In future work, we plan to allow for more regional heterogeneity of the estimation parameters—for example, by calculating different parameters for each region or subregion. Our price data also may not fully reflect the heterogeneity of prices both within and across states. Acquiring sub-state level gasoline price data or local level vehicle purchase price data could improve estimation. For example, rural households may face different gasoline and vehicle prices than do urban households, which would also affect household decisions. Nevertheless, the current results clearly demonstrate the need to account for household and regional heterogeneity in conducting policy analysis of national gasoline policies.

## References

- Archibald, R. and R. Gillingham (1981). "The Distributional Impact of Alternative Gasoline Conservation Policies." *The Bell Journal of Economics* 12(2): 426-444.
- Bento, A., L.H. Goulder, M.R. Jacobsen, and R.H. von Haefen (2008). "Distributional and Efficiency Impacts of Increased U.S. Gasoline Taxes." *American Economic Review*, 99(3): 667-699
- Berkowitz, M.K., N.T. Gallini, E.J. Miller, and R.A. Wolfe (1990). "Disaggregate Analysis of the Demand for Gasoline." *The Canadian Journal of Economics*, 23(2): 253-275.
- Berry, S., J. Levinsohn and A. Pakes (1995). "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841-890.
- . (2004). "Differentiated products demand systems from a combination of micro and macro data: The new car market." *Journal of Political Economy*, 112(1): 68-105.
- Feng, Y., D. Fullerton and L. Gan (2005). "Vehicle Choices, Miles Driven, and Pollution Policies." NBER Working Paper 11553.
- Ferguson, M., K. Ali, M. R. Olfert, and M. D. Partridge (2007). "Voting with their Feet: Jobs Versus Amenities." *Growth and Change*, 38(1): 77-110.
- Gillingham, K. (2011). "How Do Consumers Respond to Gasoline Price Shocks? Heterogeneity in Vehicle Choice and Driving Behavior." Working Paper, draft July 1, 2011.
- Goldberg, P.K. 1998. "The Effects of the Corporate Average Fuel Efficiency Standards in the US." *The Journal of Industrial Economics*, 46(1): 1-33
- Greene, D.L., and P.S. Hu. 1985. "The influence of the price of gasoline on vehicle use in multivehicle households." *Transportation Research Record*, 988: 19-24.
- Hastings, J. (2010). "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California: Reply." *American Economic Review*, 100: 1277-1279.
- Kayser, H. (2000). "Gasoline demand and car choice: estimating gasoline demand using household information." *Energy Economics*, 22: 331-348.
- Klier, T. and J. Linn (2010). The Price of Gasoline and New Vehicle Fuel Economy: Evidence from Monthly Sales Data. *American Economic Journal: Economic Policy*, 2 (3): 134-153.

- Kraft, J. and M. Rodekehr (1978). "Regional Demand for Gasoline: A Temporal Cross-Section Specification." *Journal of Regional Science* 18(1): 45-55.
- Mannering, F. and C. Winston (1985). "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization." *The RAND Journal of Economics*, 16(2): 215-236.
- Manski, C.F. (1975). "Maximum Score Estimation of the Stochastic Utility Model of Choice." *Journal of Econometrics*, 3: 205-228.
- Miller, K.K. and T.D. Rowley (2002). "Rural Poverty and Rural-Urban Income Gaps: A Troubling Snapshot of the "Prosperous" 1990s." RUPRI Data Report P2002-5.
- Partridge, M. D., D. S. Rickman, K. Ali and M. R. Olfert (2008a). "The Geographic Diversity of U.S. Nonmetropolitan Growth Dynamics: A Geographically Weighted Regression Approach." *Land Economics*, 84(2): 241-266.
- . (2008b). "Lost in space: population growth in the American hinterlands and small cities." *Journal of Economic Geography*, 8: 727-757.
- Partridge, M.D., M.R. Olfert, and K. Ali (2010). "Rural-to-Urban Commuting: Three Degrees of Integration." *Growth and Change* 41(2): 303-335.
- Petrin, A. (2002). "Quantifying the Benefits of New Products: The Case of the Minivan." NBER Working Paper 8227.
- Porter, M.E. (2003). "The Economic Performance of Regions." *Regional Studies*, 37: 549-578.
- Schmalensee, R., and T.M. Stoker (1999). "Household Gasoline Demand in the United States." *Econometrica*, 63(3): 645-662.
- Spiller, E. (2011). "Household Vehicle Bundle Choice and Gasoline Demand: A Discrete-Continuous Approach". Working paper, draft Fall 2011.
- Spiller, E., H. Stephens, C. Timmins and A. Smith (2012). "Does the Substitutability of Public Transit Affect Commuters' Response to Gasoline Price Changes?" RFF Discussion Paper.
- Storchmann, K. (2005). "Long-Run Gasoline demand for passenger cars: the role of income distribution." *Energy Economics*, 27: 25-58.
- U.S. Census Bureau (2010). Income, Poverty, and Health Insurance Coverage in the United States: 2009. <http://www.census.gov/prod/2010pubs/p60-238.pdf> (accessed December 1<sup>st</sup>, 2011).

- U.S. Energy Information Administration (EIA) (2012). How dependent are we on foreign oil?[http://www.eia.gov/energy\\_in\\_brief/foreign\\_oil\\_dependence.cfm](http://www.eia.gov/energy_in_brief/foreign_oil_dependence.cfm) (accessed July 13, 2012).
- . 2011. Motor Gasoline Sales Through Retail Outlets Prices.  
[http://www.eia.gov/dnav/pet/pet\\_pri\\_allmg\\_a\\_EPM0\\_PTC\\_Dpgal\\_m.htm](http://www.eia.gov/dnav/pet/pet_pri_allmg_a_EPM0_PTC_Dpgal_m.htm) (accessed July 7, 2011).
- U.S. Environmental Protection Agency (2012). Executive Summary: Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2010.  
<http://www.epa.gov/climatechange/emissions/usinventoryreport.html> (July 13, 2012).
- U.S. Federal Highway Administration (FHWA). 2009a. Highway Statistics 2009: Federal Tax Rates on Motor Fuels and Lubricating Oil.  
<http://www.fhwa.dot.gov/policyinformation/statistics/2009/fe101a.cfm> (accessed 7 July 2011).
- . 2009b. Highway Statistics 2009: State Motor Fuel Tax Rates, 1996–2009.  
<http://www.fhwa.dot.gov/policyinformation/statistics/2009/mf205.cfm> (accessed 7 July 2011).
- Wadud, Z., D.J. Graham and R.B. Noland (2010). "Gasoline Demand with Heterogeneity in Household Responses," *The Energy Journal*, 31(1): 47-74.
- Weber, B. and L. Jensen (2004). "Poverty and Place: A Critical Review of Rural Poverty Literature." RUPRI Rural Policy Research Center, RPRC Working Paper 04-03.
- West, S.E. (2004). "Distributional effects of alternative vehicle pollution control policies", *Journal of Public Economics*, 88: 735-757.
- . (2007). "The Effect of Gasoline Prices on the Demand for Sport Utility Vehicles", Working Paper.
- West, S.E. and R.C. Williams (2004). "Estimates from a consumer demand system: implications for the incidence of environmental taxes", *Journal of Environmental Economics and Management*, 47: 535–558
- Wu, X., J.M. Perloff, and A. Golan (2006). "Effects of government policies on urban and rural income inequality." *Review of Income and Wealth*, 52(2): 213-235.