

# Heterogeneity in the Response to Gasoline Prices: Evidence from Pennsylvania and Implications for the Rebound Effect<sup>†</sup>

Kenneth Gillingham<sup>\*</sup>, Alan Jenn<sup>\*\*</sup>, Inês Lima Azevedo<sup>\*\*</sup>

<sup>\*</sup> School of Forestry & Environmental Studies, Department of Economics, School of Management, Yale University and National Bureau of Economic Research

<sup>\*\*</sup> Department of Engineering and Public Policy, Carnegie-Mellon University

Forthcoming in *Energy Economics*

September 5, 2015

## Abstract

The consumer response to changing gasoline prices has long interested economists and policymakers, for it has important implications for the effects of gasoline taxation and vehicle energy efficiency policies. This study examines both the elasticity of driving with respect to changing gasoline prices and heterogeneity in this elasticity by geography, the fuel economy of the vehicle, and the age of the vehicle. We use detailed annual vehicle-level emissions inspection test data from Pennsylvania that include odometer readings, inspection zip codes, and extensive vehicle characteristics. We estimate a short-run gasoline price elasticity of driving demand of  $-0.10$ , and find substantial heterogeneity in this responsiveness. The elasticity is largely driven by low fuel economy vehicles, as well as vehicles between 3 and 7 years old. Our findings help reconcile some of the recent literature and provide guidance on the magnitude of the direct rebound effect from light duty vehicle energy efficiency policies.

JEL classification codes: H23, Q38, Q41.

Keywords: energy efficiency, rebound effect, gasoline price elasticity, big data analytics.

---

<sup>†</sup> We thank audiences at the 6<sup>th</sup> Atlantic Workshop on Energy & Environmental Economics, the Industry Studies Association 2014, and the 2013 Behavior, Energy, and Climate Change Conference, as well as Jeremy Michalek, Pedro Ferreira, Paul Fischbeck, and Scott Matthews for comments. This work was supported by the center for Climate and Energy Decision Making (SES-0949710), through a cooperative agreement between the National Science Foundation and Carnegie Mellon University. We also acknowledge support from a Bertucci Fellowship through Carnegie Mellon University. Corresponding author: Kenneth Gillingham, Yale University and NBER, 195 Prospect Street, New Haven, CT 06510, 203-436-5465, kenneth.gillingham@yale.edu.

# 1 Introduction

The United States consumes over 130 billion gallons of gasoline annually, with the vast majority used in the transportation sector. The use of fossil fuel in transportation leads to nearly 2 billion metric tons of CO<sub>2</sub> being emitted into the atmosphere every year in the US—approximately one-third of US greenhouse gas (GHG) emissions (EPA, 2014). In an effort to reduce these emissions and imported oil, policymakers have implemented a variety of measures, from Corporate Average Fuel Economy (CAFE) standards and the Zero Emissions Vehicle (ZEV) mandate, to Renewable Fuel Standards (RFS) and Low Carbon Fuel Standards (LCFS). In addition, the gasoline tax has been falling in real terms across the United States (Jenn et al., 2015). The effects of changes in policies such as gasoline taxes and fuel economy standards depend critically on the consumer responsiveness to changes in relative prices engendered by the policies.

This paper examines how consumers change the amount they drive in response to changing gasoline prices—and the heterogeneity in this response. Such heterogeneity could come about due to different preferences, different urban spatial structures, and differences in the cost or appeal of driving different types of vehicles. In this study we address the following questions: are some consumers more responsive than others? Is the response greater for different types of vehicles than others, perhaps by fuel economy or age of vehicle? To answer these questions, we use a massive dataset of over 75 million inspection records, which include odometer readings, covering 2000 to 2010.

We find a point estimate for the one-year gasoline price elasticity of driving of -0.10. However, quantile regressions reveal that a high percentage of the vehicles are almost entirely inelastic in response to gasoline price changes. Further, we find that the lowest fuel economy vehicles in the fleet drive the responsiveness, with higher fuel economy vehicles highly inelastic with respect to gasoline price changes. Similarly, we find considerable heterogeneity in the response by the age of the vehicle, with drivers of vehicles in the age bracket of 3 to 7 years being the most responsive.

These results not only shed light on the differing effects of changes in gasoline prices, but also have clear implications for policy. If consumers respond in the same way to changes in gasoline prices and gasoline taxes, as might be expected, then our estimates can be interpreted as the effect on driving from a marginal increase in the gasoline tax.<sup>1</sup> Our results thus provide guidance for the additional driving that might be expected due to declining real gasoline taxes.

In addition, they also have implications for the gasoline and greenhouse gas savings possible from fuel economy standards. With an improvement in fuel economy, the cost per mile of driving decreases and consumers would be expected to drive more. This effect has come to be known in the energy economics literature as the direct rebound effect: an increase in energy use that reduces or even eliminates the energy and greenhouse gas savings that would have been expected from the efficiency improvement (Azevedo, 2014; Gillingham et al., 2016; Sorrell, 2007; Thomas and Azevedo, 2014). It is considered a key parameter in the analysis of the costs and benefits of fuel economy standards; indeed findings of positive net benefits often turn on this parameter (Gillingham et al., 2016).

It is common convention to use the price elasticity of demand for driving as a proxy for direct rebound effect, a convention that is valid under several assumptions, including that there are few substitute goods and that consumers respond the same way to changes in the cost per mile of driving regardless of the source of the variation (e.g., gasoline price changes or fuel efficiency changes) (Chan and Gillingham, 2015; Thomas and Azevedo, 2013).<sup>2</sup> Thus, our baseline result may be interpreted as a proxy for the rebound effect, while the finding of heterogeneity in responsiveness by vehicle type may extend to heterogeneity in the rebound effect. However, as the true direct rebound effect is by definition calculated based on an unobserved counterfactual where fuel economy (not fuel prices) change, we are cautious in this interpretation. Moreover, our findings of heterogeneity in the response to gasoline prices by

---

<sup>1</sup> It is possible that consumers respond differently to changes in taxes than to changes in prices, perhaps due to differences in salience (Li et al., 2014).

<sup>2</sup> An important reason for using gasoline price variation is that there is usually only cross-sectional variation in fuel economy, which is less plausibly exogenous and perhaps more difficult to instrument for than time series variation in gasoline prices.

vehicle fuel economy and age raise further questions about the validity of the assumption that consumers respond the same way to changes in the cost per mile of driving regardless of the source of the variation.

The remainder of our paper is structured as follows. The next section discusses the relevant literature, while Section 3 describes our unique and rich registration dataset. Section 4 covers our empirical strategy for estimating driving demand. Section 5 presents our results and Section 6 concludes.

## 2 Relevant Literature

This paper relates to an extensive empirical literature on gasoline price elasticities of demand, but is one of only a few papers to estimate a gasoline price elasticity of demand for driving, rather than demand for gasoline. Of course, the two elasticities are closely related, for the price elasticity of gasoline demand can be decomposed into responses on both the intensive margin (i.e., driving) and extensive margins (i.e., new vehicle purchase and scrappage), with the gasoline price elasticity of driving capturing the response on the intensive margin (Gillingham, 2011). One of the earliest reviews of studies estimating the price elasticity of the demand for gasoline finds a mean (across studies) short-run estimate of -0.26 based on studies from 1979 to 1991 (Dahl and Sterner, 1991). A somewhat more recent review of studies from 1967 to 2000 finds estimates in the range of -0.09 to -0.31 (Greening et al., 2000).<sup>3</sup> These earlier reviewed estimates did not have the benefit of the rich datasets available today and in many cases did not carefully consider possible endogeneity concerns.

More recently, there has been work indicating that the price elasticity of gasoline demand has been decreasing over time in the United States. Using different approaches and datasets, both Small and Van Dender (2007b) (using pooled cross section data for the US states from 1966 to 2001) and Hughes et al. (2008) (who examined gasoline demand from 1975 to 1980 and 2001 to

---

<sup>3</sup> Other notable reviews include Goodwin et al. (2004) and Graham and Glaister (2004), both which review estimates ranging widely from near-zero to close to -1.

2006) show that during the period of relatively low and stable gasoline prices in the late 1990s and early-to-mid 2000s, consumers appear to have become even more inelastic, with estimated short-run gasoline demand elasticities in the -0.03 to -0.08 range. These findings suggest that consumer responsiveness may vary depending on conditions, much as Gately and Huntington (2002) show that there is an asymmetric consumer response to increases and decreases in oil prices.

The possibility of varying consumer responsiveness is further underscored by a set of more recent results that include data from the late 2000s, when gasoline prices were rapidly changing. Hymel and Small (2014) update the previous Small and Van Dender (2007b) analysis to 2009, and find evidence of a slightly greater responsiveness during this time period. This result would help reconcile the previous findings with recent results using odometer reading data in the late 2000s showing a much more elastic response in the gasoline price elasticity of *driving* in California; Gillingham et al. (2013) estimates a “two-year” gasoline price elasticity of driving for new vehicles of -0.22 and Knittel and Sandler (2013) estimates a “two-year” gasoline price elasticity of driving for all but the newest vehicles of -0.15. This difference already suggests that consumers of vehicles of different ages may respond differently, possibly due to different needs, preferences, or locations.

Our finding of a greater responsiveness for drivers of lower fuel economy vehicles corresponds with similar findings from California by Knittel and Sandler (2013) and Germany by Frondel et al. (2012). What is new about our result is that it is based on actual annual odometer reading data from all vehicles in a large area, rather than just older vehicles in California or trip diary data from Germany. Similarly, Gillingham (2011), Frondel et al. (2012) and Gillingham (2014) also use quantile regression approaches to examine the driving responsiveness to gasoline prices, but this is the first paper to both examine all vehicles using revealed odometer readings and include fixed effects in the quantile regressions to address potentially correlated unobservables.

A few key points emerge from the literature, which are relevant to interpreting our results. Since these estimated elasticities are often being used in policymaking as a proxy for the rebound

effect from fuel economy standards, they suggest that the rebound effect may be changing over time and across geography. Furthermore, the few studies that attempt to estimate the rebound effect more directly through an elasticity of vehicle-miles-traveled (VMT) with respect to fuel economy, such as Small and Van Dender (2007b), generally are in the same order of magnitude, and with similar variation, as the studies discussed above that focus on the fuel price elasticity (Azevedo, 2014). While we are careful about interpreting our estimates directly as a rebound effect, we believe that they contribute to the literature suggesting that the price elasticity of driving and the rebound effect may be heterogeneous and relatively close to zero in the short run.

### **3 Data**

#### **3.1 The Pennsylvania Inspection Program**

This paper uses a dataset of over 7 million light duty vehicles, covering the years 2000 to 2010, resulting in over 75 million vehicle inspection records from mandatory annual inspections in Pennsylvania. These annual inspections involve both a visual safety check and an on-board diagnostic (OBD) test that electronically records and submits emissions, odometer readings, zip code of the inspection station, other vehicle information, and the date of inspection to the Pennsylvania Department of Transportation. All light duty vehicles in Pennsylvania with a model year after 1975, including hybrid electric vehicles, are required to undergo the mandatory safety and emissions inspection. In addition, a new inspection is required any time there is an update to a vehicle registration, such as a change of address or ownership. In the case of a failed inspection, drivers must repair the vehicle and follow up with a free inspection within 30 days.<sup>4</sup>

The mandatory inspection program was rolled out over a several year period and with start dates differing across Pennsylvania regions. The roll out of the inspection program started with the most populated regions (e.g., the Philadelphia and Pittsburgh metropolitan areas) in 2000. These regions have the greatest air quality challenges and thus were prioritized for the program. In the following years, the program was rolled out to other regions that also faced air quality challenges.

---

<sup>4</sup> Drivers may also opt to receive a 1-year waiver if repair costs exceed \$150.

Figure 1 illustrates the roll-out of the inspection program with a map of Pennsylvania that shows when mandatory inspections began in each zip code. This figure illustrates that all of the most populated areas in Pennsylvania are required to undergo a mandatory inspection by around 2005. The areas of the state that do not require mandatory inspections tend to be rural and mountainous, and host a smaller share of the Pennsylvania light duty vehicle fleet.<sup>5</sup>

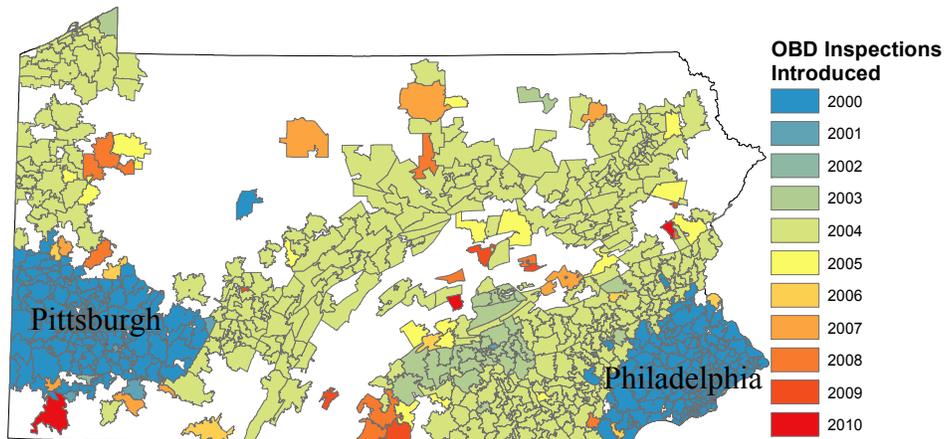


Figure 1: Zip codes by date for which the inspections requirements were introduced in Pennsylvania, from 2000 through 2010. Annual inspections are not required in the white areas of the state. Source: <http://www.drivecleanpa.state.pa.us/>.

The inspection program only covers light duty passenger vehicles, which make up roughly 70 percent of the vehicle fleet in Pennsylvania (PennDOT, 2010). Figure 2 shows the percentage of the total light duty vehicle fleet in Pennsylvania covered by the inspection program in each year. By the end of our sample, in 2010, roughly two-thirds of all light duty vehicles undertook an inspection and are included in our dataset. Any vehicles not recorded were either in violation of the law, in zip codes that were not subject to the mandatory inspections or were sold before 1975. Since the program roll out took some time to get started in late 2000, the first inspections for most vehicles in Philadelphia and Pittsburgh were in 2001. The same is true for vehicles in regions where the program was rolled out in 2004; these vehicles generally had their first inspections in 2005.

<sup>5</sup> A full list of exempted regions can be found at <http://www.drivecleanpa.state.pa.us/map.htm>.

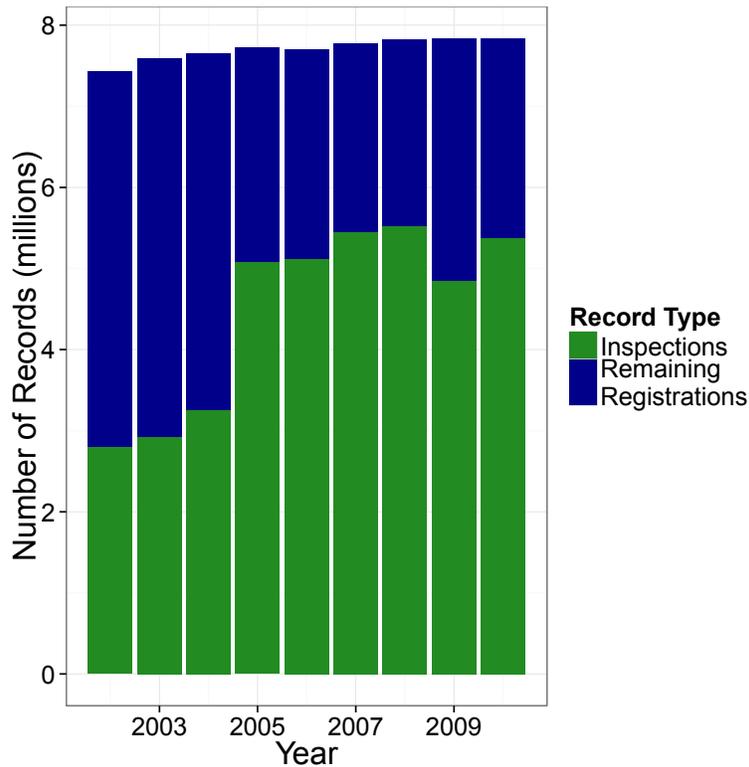


Figure 2: Total registered vehicles (blue + green) and number of inspection records in our dataset (green).

Sources: PennDOT inspection data and

<http://www.dmv.state.pa.us/stats/pennsylvaniaMotorVehicleRegs.shtml>.

In the regions covered by the inspection program, there are important patterns in the data that bear relevance to interpreting our results. First, vehicles are not located uniformly. Certain regions of Pennsylvania make up a much larger portion of our dataset than others, as is illustrated in Figure 3, which plots the number of inspections by the zip code of the inspection. Notably, the areas with the greatest number of inspections appear to be in the larger Philadelphia and Pittsburgh metropolitan areas (largely in the suburbs of these cities). Thus, these areas will be heavily weighted in our regression results.

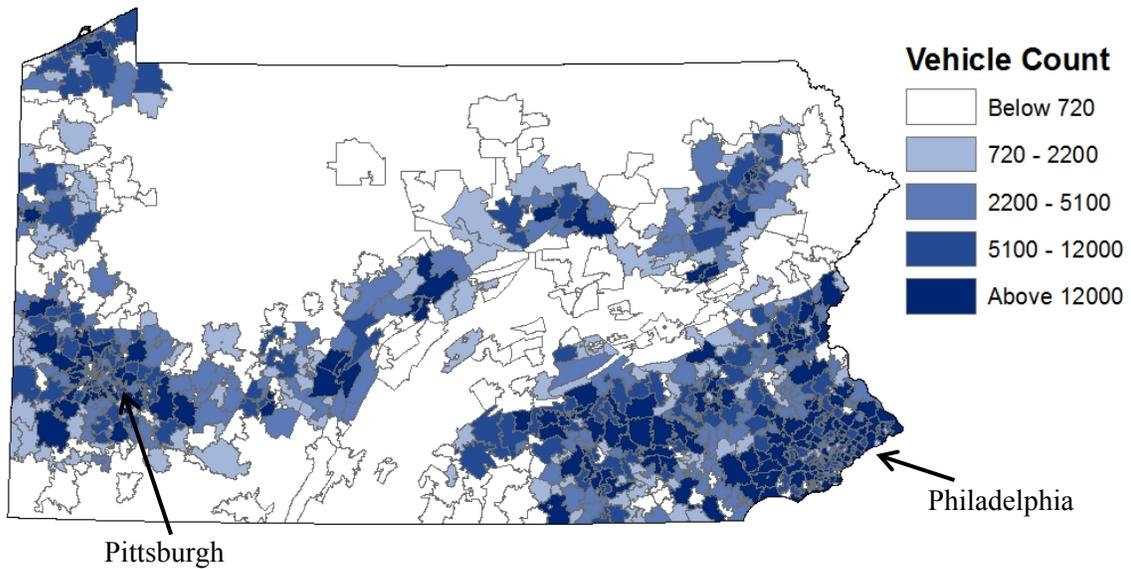


Figure 3: Number of vehicles in our dataset by zip code in 2010. Source: PennDOT inspection data.

Second, the areas with the most vehicle inspections per household are distributed across Pennsylvania in a somewhat more uniform pattern (Figure 4). In general it appears that the suburban areas in the Philadelphia and Pittsburgh metropolitan areas have slightly more vehicles per household than other areas, but there is no strong and obvious pattern. One caveat to this is that we only observe the zip code where vehicles are inspected, which may differ slightly from where the vehicles are garaged.

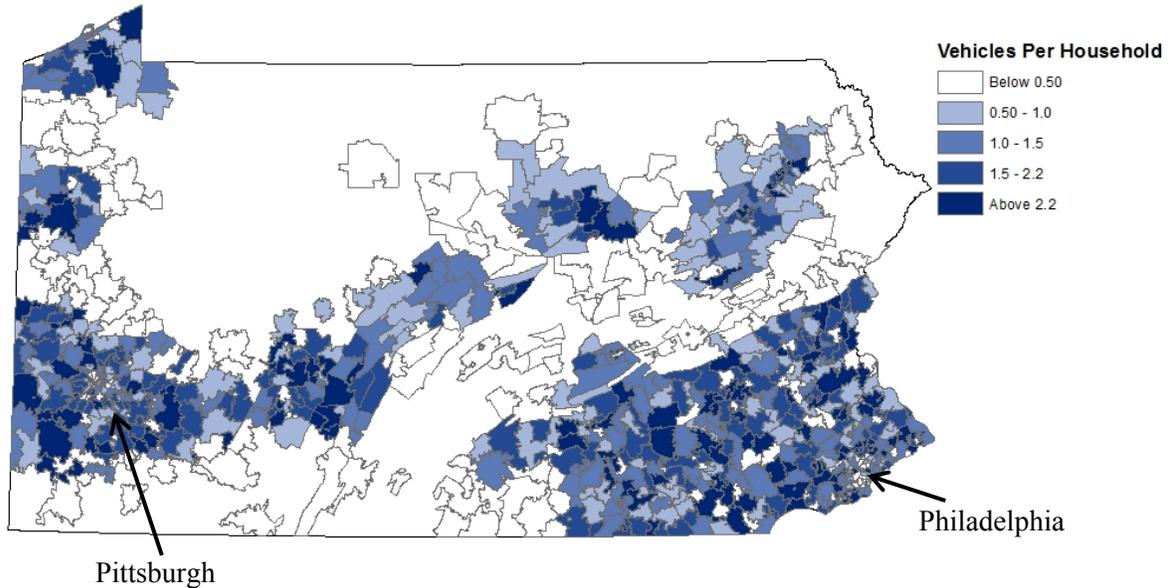


Figure 4. Vehicles per household by zip code in Pennsylvania in 2010. Source: PennDOT inspection data and households from the 2010 US Census.

### 3.2 Construction of the Panel Dataset

We begin with over 75 million vehicle inspection records over the period 2000 to 2010, each with an odometer reading. We are interested in the miles driven during the period between two inspections, and follow the strategy from several other recent papers in the literature that have odometer reading data (Gillingham, 2014; Knittel and Sandler, 2013) by using the vehicle-driving period as our unit of observation. Consider a case where one inspection was on July 1, 2001, with an odometer reading of 30,000, and the next inspection was on June 1, 2002, with an odometer reading of 41,000. Then a vehicle-driving period level observation for this driving period would have a VMT of 11,000, a time between inspections of 11 months, and a second test date of June 1, 2002. For interpretability, we normalize VMT to be in terms of VMT per year based on the number of days in the driving period. So in our simplified example, the VMT per

year would be coded to 12,000. Converting our data to the vehicle-driving period level leaves us with 34,596,513 observations.<sup>6</sup>

Each vehicle in the dataset is indexed by a Vehicle Identification Number (VIN), which we decode using a VIN decoder from DataOne, Inc. to reveal the following vehicle information: make, model, model year, body style, body type, engine type, engine size, fuel type, drive type, highway fuel economy in miles per gallon (MPG), weight, and manufacturer's suggested retail price (MSRP). We create a variable for vehicle age by subtracting the model year from the year of the inspection.<sup>7</sup> We are able to decode the VINs for 97 percent of the observations in the dataset. The remainder are either invalid VINs or for vehicles prior to 1980 and we drop these observations.

Since we are interested in the response in VMT to changing gasoline prices, we bring in monthly Pennsylvania gasoline price data from EIA (2014). Similarly, we bring in macroeconomic conditions: the monthly average U.S. gross domestic product (GDP) from the Bureau of Economic Analysis (BEA, 2014) and the average unemployment rate in Pennsylvania based on data from the Bureau of Labor Statistics (BLS, 2014). Finally, we also bring in data on major refinery supply disruptions in the Gulf Coast (primarily due to hurricanes) from the Energy Information Administration (EIA, 2011). All four of these variables are at the monthly level in raw data form and are converted to the vehicle-driving period level. This is accomplished by taking the average value for each over the months of the driving period covered by each observation. This means that for vehicle  $i$  that performed an inspection in time  $t$  and the subsequent inspection in time  $t + m$ , we compute the average gasoline price, GDP, unemployment, and supply disruptions between  $t$  and  $t + m$  for that observation. So, for example, we compute the price of gasoline for the driving period of a vehicle inspected first on July 1st,

---

<sup>6</sup> The 34.6 million driving periods are less than half of the 75 million inspection records because of vehicles that only had one inspection, such as new vehicles or vehicles that left the state shortly after the program began.

<sup>7</sup> We recognize that some vehicles may be sold before the model year or after the model year, so this is a proxy for the actual vehicle age. However, this proxy is what is typically used for used car pricing (e.g., see the Kelley Blue Book), so it is as good as possible without actual data on when the new vehicle was first purchased.

2001 and then on June 1st, 2002 as the average of the monthly gasoline prices in the 11 months between inspections.

Finally, to address potential outliers, we truncate our data and drop observations where vehicles were not used (e.g. a VMT of 0), as well as vehicles with VMT per driving period greater than 100,000 miles or odometer readings greater than 7,000,000 miles, which are likely inspection reporting errors. The final panel dataset contains 30,621,721 observations and 7,173,110 distinct vehicles. We convert all dollar-valued variables to real 2010 dollars using the BLS Consumer Price Index.

### 3.3 Data Description

Table 1 presents summary statistics for key variables in the final full dataset. The average annual VMT is 10,183 miles, with a standard deviation that is nearly as large, at 7,303 miles. The average gasoline price is \$2.46/gallon, with a standard deviation of \$0.55/gallon.

Table 1: Summary statistics for selected variables in the full dataset.

	Mean	Std. dev.	Min.	Max.
Annual VMT (mi)	10,183	7303	0.137	100,000
Average Gas Price (2010\$/gal)	2.46	0.55	1.23	4.14
MPG Highway (mi/gal)	20.1	4.20	8.46	58.14
Age (years)	7.64	4.59	0.00	29.00
Avg Adjusted Unemployment (%)	5.45	0.97	4.23	8.62
Avg GDP (billions of 2010\$)	14,015	710	12,634	15,056
Avg Refinery Disruptions (millions of barrels)	2.52	3.42	0.001	70.48

Notes: All variables have 30,621,721 non-missing observations. The age of the vehicle is in the number of years between the model year and the year of the second test in a driving period.

Figure 5 illustrates the variation in annual VMT in Pennsylvania over the full 2000-2010 sample. The VMT distribution appears to have a zero-truncated normal distribution, with a long

tail of drivers who drive a substantial amount (e.g., 60,000 miles per year is over 160 miles per day).

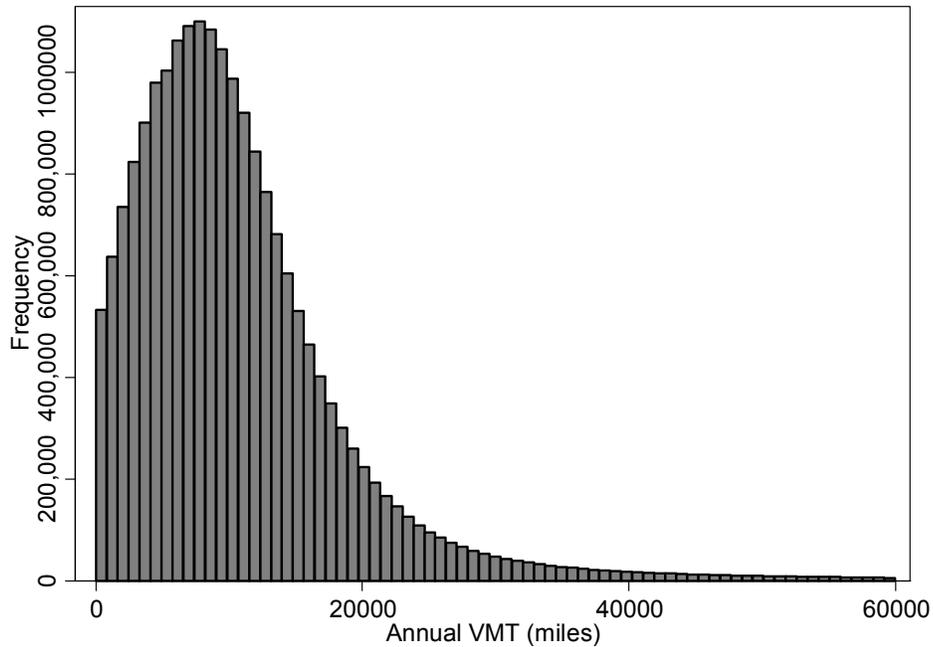


Figure 5. Histogram of the annual vehicle miles driven in Pennsylvania. Source: PennDOT inspection data.

The variation in driving may stem from a variety of sources, including geographic heterogeneity. Figure 6 shows how VMT differs considerably across Pennsylvania (for observations from all years), with vehicles being driven the most located in rural and suburban areas. This likely reflects the greater need to drive in these areas.

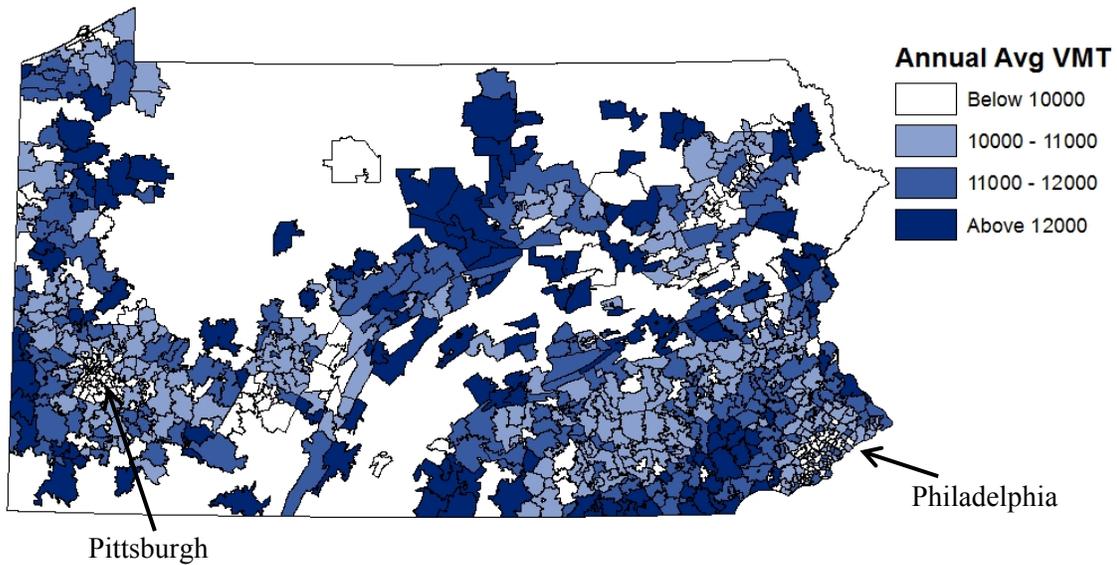


Figure 6. Average annual VMT by zip code, 2000-2010. Source: PennDOT inspection data.

Figure 7 shows the changes in the average gasoline price and macroeconomic conditions over the time period of analysis. Gasoline prices display considerable time-series variation, with a dramatic and sustained increase in prices from 2004 to 2008, followed by a steep decline in 2008 and a leveling off afterwards. This variation will be leveraged for identifying our coefficients.

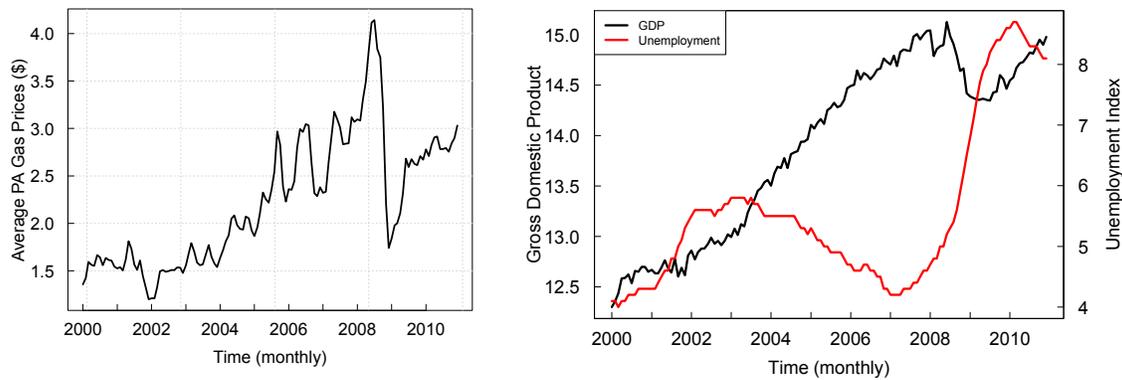


Figure 7. Gasoline prices (in \$2010) have considerable time series variation during our time period (Left). Macroeconomic conditions (\$2010 trillion U.S. GDP and Pennsylvania unemployment) also changed substantially over this time (Right). Sources: US EIA, US BEA, US BLS.

In Figure 8, we provide descriptive evidence on trends in the average annual gasoline price and average annual VMT over each driving period in our full sample. The solid lines show that the mean VMT has decreased slightly over this period, while the mean gasoline price has increased over this period. The wide bands of the dotted around the solid lines show plus or minus one standard deviation in the variables (based on the year of the second inspection of the driving period).

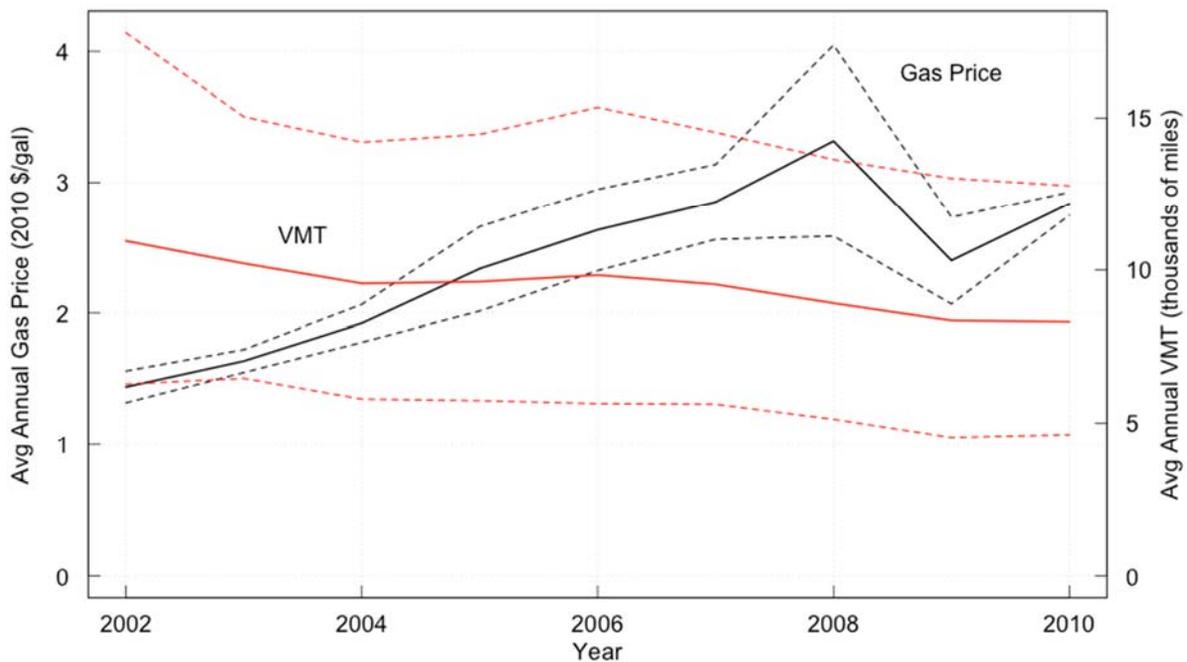


Figure 8. The average annual gasoline price and VMT are shown in the solid lines, with +/- one standard deviation shown in the dotted lines for each. The year on the x-axis denotes the year of the second test of a driving period. Sources: US EIA and PennDOT inspection data.

Figure 8 descriptively suggests only a relatively minor response in driving to gasoline prices. Of course, as shown in Figure 7, macroeconomic conditions also changed, so our empirical specification will control for these changes.

As mentioned above, vehicles in the Philadelphia and Pittsburgh metropolitan areas not only have more vehicles, but also began receiving inspections much earlier than the rest of Pennsylvania, most of which began inspections after 2004. To illustrate the differences between these two subsamples, Table 2 provides summary statistics for the subsample of all vehicles that begin inspections in 2000 (Philadelphia and Pittsburgh) and all vehicles that began inspections in 2004 (many other parts of the state). Note that both the time frames and the regions differ, so one must be cautious in interpreting these as only regional differences or differences due to time.

Table 2: Summary statistics by geographic region based on introduction of OBD inspections.

	Mean	Std dev.	Min.	Max.
<b><i>OBD Inspections begun in 2000</i></b>				
Annual VMT (mi)	9,951	7342	0.138	100,000
Average Gas Price (2010\$/gal)	2.25	0.57	1.23	4.14
Combined MPG (mi/gal)	20.1	4.09	8.80	58.14
Age (years)	8.07	4.43	0.00	29.00
Avg Adjusted Unemployment (%)	5.45	0.97	4.23	8.62
Avg GDP (trillions of 2010\$)	14,015	710	12,634	15,056
Avg Refinery Disruptions (millions of barrels)	2.64	3.64	0.001	70.48
<b><i>OBD Inspections begun in 2004</i></b>				
Annual VMT (mi)	9,861	6,376	0.31	100,000
Average Gas Price (2010\$/gal)	2.77	0.55	1.23	4.14
Combined MPG (mi/gal)	19.8	4.27	8.80	58.14
Age (years)	7.48	4.35	0.00	29.00
Avg Adjusted Unemployment (%)	5.78	1.55	4.23	8.62
Avg GDP (trillions of 2010\$)	14,640	197	13,256	15,056
Avg Refinery Disruptions (millions of barrels)	2.62	3.37	0.007	60.47

Notes: All variables in the 2000 inspections panel have 18,642,460 non-missing values, while all variables in the 2004 inspections panel have 3,097,592 non-missing values. Note that these do not add up to 30 million because the remaining vehicles entered the dataset in other years. The age of the vehicle is in the number of years between the vehicle model year and the year of the second test in a driving period.

## 4 Empirical Approach

### 4.1 Model specification

We model the demand for driving vehicle  $i$  with a test in year  $t$  and  $j$  months between tests as a function of the Pennsylvania gasoline price ( $P_{it}$ ), national GDP ( $G_{it}$ ), unemployment rate in Pennsylvania ( $U_{it}$ ) the age of the vehicle ( $A_{it}$ ), vehicle characteristics ( $\theta_i$ ), other broader time-varying factors ( $\eta_t$ ), month-of-the-year seasonal factors ( $\zeta_s$ ) and any factors that relate to the length of the driving periods ( $\mu_d$ ):

$$VMT_{it} = f(P_{it}, G_{it}, U_{it}, A_{it}, \theta_i, \eta_t, \zeta_s, \mu_d).$$

Controls for factors that relate to the length of the driving period are included because there may be potentially correlated unobserved heterogeneity in drivers based on when they receive their inspection (e.g., those who get a late inspection may drive differently than those who get an on-time inspection, and their driving period will cover a different time frame of gasoline prices).

We further specify demand as follows:

$$VMT_{it} = (P_{it})^\gamma (G_{it})^{\beta_G} (U_{it})^{\beta_U} \exp(\beta_0 + g(A_{it}) + \theta_i + \eta_t + \zeta_s + \mu_d + \varepsilon_{it}),$$

where  $g(\cdot)$  is a quadratic function in age, which flexibly captures the relationship between vehicle age and driving.  $\varepsilon_{it}$  is a mean-zero stochastic error term. This specification implies the following log-log form:

$$\log(VMT_{it}) = \beta_0 + \gamma \log(P_{it}) + \sum_{X \in M} \beta_M \log(X) + g(A_{it}) + \theta_i + \eta_t + \zeta_s + \mu_d + \varepsilon_{it}, \quad (1)$$

where  $M = \{G_{it}, U_{it}\}$  and  $\gamma$  is our parameter of interest, which can be interpreted as the elasticity of VMT with respect to the price of gasoline. Given that inspections are generally once a year, any estimate is best understood as a one-year elasticity of driving (i.e., a short-run elasticity).  $\theta_i$  are vehicle fixed effects.  $\eta_t$ ,  $\zeta_s$ , and  $\mu_m$  are indicator variables for the year that the driving period

finishes, the month-of-the-year that the driving period finishes, and driving period length categories (early tests that are less than 11 months, normal tests between 11 months and 13 months, late tests between 14 and 21 months, and very late tests beyond this).<sup>8</sup>

We estimate this equation using a standard fixed effects regression as well as using a quantile regression (Koenker and Bassett, 1978) to understand the heterogeneity in responsiveness across consumers. Furthermore, to better understand the sources of heterogeneity, we also estimate (1) using partitions of our dataset based on fuel economy and vehicle age. The model in (1) follows most of the existing literature in that it implies that consumers respond symmetrically to increases and decreases in gasoline prices. In this sense, our results—just like most others—should be viewed as estimating an average response.

## 4.2 Identification and potential estimation concerns

With the inclusion of such a rich set of fixed effects, the model specifications we estimate address a variety of potential empirical concerns. For example, vehicle fixed effects capture unobserved heterogeneity in time-invariant vehicle characteristics, location, and driver attributes.<sup>9</sup> Gillingham (2013) shows that estimates of the demand for driving can be biased upwards due to unobserved preferences for driving that lead consumers to purchase a more efficient vehicle. By including vehicle fixed effects, we address this unobserved heterogeneity in a flexible non-parametric way. Year-of-test indicator variables address time-varying factors besides the macroeconomic conditions that we already control for. Month-of-the-year indicator variables capture the possibility that different types of consumers purchase vehicles at different times of year (Copeland et al., 2011), and thus have inspections performed during that time of year in subsequent years. Finally, indicator variables for the length of the driving period flexibly control for possible heterogeneity in consumers who have their inspection done on time or at different times.

---

<sup>8</sup> Our results are robust to the exact specification of these indicator variables.

<sup>9</sup> Note that some vehicles move the location of their inspections during our time period. These movers correspond to much less than 1% of the observations in the dataset. We observe these movers and find that excluding the movers from our analysis makes little difference. Thus we opt to include them and simply use vehicle fixed effects.

Including the rich set of fixed effects is very useful for identifying the relationship between gasoline prices and VMT by addressing a variety of potential confounders. Yet, we recognize that there remains a possible concern about the classic simultaneity of supply and demand. This issue is beautifully described on page 1 in Kennan (1989):

“For example, even if each buyer has a negligible influence on the market price, there may still be an important correlation between the price and the disturbance in an individual demand equation, so that treating price as econometrically exogenous gives biased coefficient estimates.”

Fundamentally, the concern is that unobserved local demand shocks will shift demand, increasing the gasoline price in equilibrium and leading to a supply response (e.g., immediately refining more gasoline and moving it to Pennsylvania). We would then be estimating an equilibrium response, rather than actually pinning down the demand curve.

While we expect that our rich set of fixed effects should address most of these issues, we grant that some may remain. Thus, in our preferred specification, we instrument for the gasoline price with an exogenous supply shifter: major gasoline refining supply disruptions in Gulf Coast due to hurricane activity. This instrument is highly correlated with the Pennsylvania gasoline price, due to the integration of the gasoline market in much of the United States. Yet there is no reason to believe that it is correlated with local demand factors in Pennsylvania, except through the gasoline price. A similar oil price shock instrument was also used in Hughes et al. (2008).

It is worth noting that our empirical strategy, while flexibly controlling for key estimation concerns, does impose some structure to the relationship between gasoline prices and VMT. There has been a series of papers on the gasoline price elasticity of demand that use semi-parametric or non-parametric estimation approaches (Blundell et al., 2012; Hausman and Newey, 1995; Liu, 2014; Schmalenensee and Stoker, 1999; Wadud et al., 2010b; Yatchew and No, 2001). While of merit, these approaches often also impose some structure, albeit less than in our specification. But more importantly these approaches cannot flexibly control for potential confounders with a rich set of fixed effects (usually because the data available for estimation

does not permit rich sets of fixed effects). We deem that the value of the fixed effects approach, and especially vehicle fixed effects, for identification of our coefficients is sufficient to opt for our approach rather than a semi-parametric or non-parametric approach.<sup>10</sup>

## 5 Results

### 5.1 The gasoline price elasticity of driving

We begin by estimating (1) on the full dataset of over 30 million observations. Table 3 shows our primary results. Column 1 contains year ( $\eta_t$ ), month-of-the-year ( $\zeta_s$ ), and driving period length ( $\mu_d$ ) indicator variables. Column 2 contains vehicle fixed effects ( $\theta_i$ ), but none of the other indicator variables. Column 3 contains year, month-of-the-year, and months-to-test indicator variables, as well as vehicle fixed effects. Column 4 is the same as column 3, but with the log of the average gasoline price instrumented for using the log of the average size of the oil refinery disruptions (in millions of barrels of oil). The first-stage results from this IV estimation are available in the appendix. The instrument is very strong: the  $t$ -statistic on the log refinery disruptions coefficient in the first-stage is over 1,000.

---

<sup>10</sup> Gillingham and Munk-Nielsen (2015) explore this issue in more detail using rich data from Denmark and find similar results with semi-parametric and vehicle fixed effects approaches. We also explore other functional forms, such as a linear form, and find largely similar results.

Table 3: Primary regression results with full dataset.

	(1)	(2)	(3)	(4)
	Time Controls	Vehicle FE	All FE	All FE; IV
ln(gasoline price)	-0.143*** (0.004)	-0.132*** (0.003)	-0.219*** (0.003)	-0.099*** (0.008)
ln(GDP)	6.32*** (0.033)	2.09*** (0.021)	3.85*** (0.041)	3.18*** (0.056)
ln(unemployment)	0.765*** (0.004)	0.213*** (0.002)	0.311*** (0.005)	0.328*** (0.004)
age	-0.011*** (2.21E-4)	-0.059*** (4.19E-4)	-0.008*** (8.40E-4)	-0.091*** (7.38E-4)
age <sup>2</sup>	-0.003*** (1.24E-5)	-0.002*** (1.66E-5)	-0.002*** (1.66E-5)	-1.79E-3*** (1.52E-5)
vehicle fixed effects		X	X	X
year dummies	X		X	X
month-of-year dummies	X		X	X
driving period dummies	X		X	X
R-squared	0.18	0.71	0.71	0.05
N	30,621,721	30,621,721	30,621,721	30,621,721

Notes: Dependent variable is the log average VMT over a driving period. Column 4 instruments for the log gasoline price with the log of the average size of the refinery disruptions over the driving period. For column 4, the within R-squared is reported. Robust standard errors clustered on vehicle in parentheses.

\*\*\* denotes significant at the 1% level.

In all four specifications, the coefficient on the log of the average gasoline price—our estimated gasoline price elasticity of driving—indicates a highly inelastic response to gasoline prices. The coefficient ranges from -0.10 to -0.21. Our preferred specification is in column 4, which implies a one-year price elasticity of demand for driving of -0.10. This result of an inelastic short-run response to gasoline prices is in line with much of the literature for the United States. For context, it implies slightly more responsiveness than Small and van Dender (2007a)

and slightly less responsiveness than estimates by Gillingham (2014) and Knittel and Sandler (2013) for California. It also implies much less responsiveness than estimates using German travel diary data by Frondel et al. (2012), which show elasticities in the range of -0.4 to -0.7. These differences across studies over different times and locations may belie the heterogeneity in responsiveness by time frame and location.

Our other coefficients in Table 3 are included as controls. The macroeconomic variables are positive and statistically significant in all specifications. There is a sensible interpretation for the coefficient on GDP: more economic growth implies more driving. The positive coefficient on unemployment may indicate that when unemployment is high workers are more likely to take jobs further away from home, leading to more driving. This especially makes sense after controlling for GDP, since GDP most likely captures changes in leisure driving due to increased income that could possibly correlated with gasoline prices.<sup>11</sup>

The vehicle age coefficients suggest that vehicles are driven less as they age, consistent with previous findings in other settings (Gillingham, 2011). The quadratic term is very close to zero, so the relationship appears to be close to linear. The relationship is slightly concave, consistent with driving a similar amount for several years as a vehicle ages and then dropping more steeply in driving once the vehicle is older, as one might expect. Older vehicles are often kept as a second or third vehicle in a household and they may be driven by retired drivers. We cannot test these possible explanations in our data. However, we do explore the responsiveness by the age the vehicle in section 5.4.

One interesting feature of our dataset is that in the first few years of our sample we only observe vehicles in the Philadelphia and Pittsburgh metropolitan areas, while mid-way through the sample, we begin to observe vehicles from many other areas in Pennsylvania (Figure 2). While our year fixed effects should largely address any concern from the sample changing over time, we also perform the estimations separately on the Philadelphia and Pittsburgh metropolitan

---

<sup>11</sup> Removing unemployment altogether from the regression does not change our coefficient on the gasoline price in any substantive way.

areas and on the remaining areas in Pennsylvania to provide insight into how the response may differ.

Table 4 shows the results of these estimations of Equation (1) on the two subsamples. Columns 1 and 2 estimate the model for the metropolitan areas in Pennsylvania (from 2000 to 2010), while columns 3 and 4 estimate the model on the remaining areas of Pennsylvania that require an inspection starting in 2004 (from 2004 to 2010). All columns have the full set of fixed effects and indicator variables, while columns 2 and 4 instrument for the gasoline price.

Table 4: Estimation on the two subsamples based on year of mandatory OBD inspections.

	(1)	(2)	(3)	(4)
	OBD 2000	OBD 2000; IV	OBD 2004	OBD 2004; IV
ln(gasoline price)	-0.066*** (0.063)	-0.076*** (0.015)	-0.193*** (0.011)	-0.150*** (0.0237)
ln(GDP)	2.31*** (0.053)	3.33*** (0.082)	4.38*** (0.164)	4.31*** (0.240)
ln(unemployment)	0.301*** (0.006)	0.398*** (0.006)	0.412*** (0.016)	0.429*** (0.014)
age	-0.0763*** (0.001)	-0.0973*** (9.25E-4)	-0.0852*** (0.003)	-0.0891*** (0.002)
age <sup>2</sup>	-1.69E-3*** (2.07E-5)	-1.76E-3*** (1.84E-5)	-2.46E-3*** (4.49E-5)	-2.42E-3*** (3.98E-5)
vehicle fixed effects	X	X	X	X
year dummies	X	X	X	X
month-of-year dummies	X	X	X	X
driving period dummies	X	X	X	X
R-squared	0.62	0.05	0.76	0.05
N	18,642,460	18,642,460	3,097,592	3,097,592

Notes: Dependent variable is the log average VMT over a driving period. Columns 1 and 2 are estimated on the subsample that first required inspections in 2000 (i.e., Philadelphia and Pittsburgh metro areas). Columns 3 and 4 are estimated on the subsample that first required inspections in 2004. Columns 2 and 4 instrument for the log

---

gasoline price with the log of the average size of the refinery disruptions over the driving period. The within R-squared is presented for columns 2 and 4. Robust standard errors in parentheses, clustered at the vehicle level. \*\*\* denotes significant at the 1% level.

We find that our results are relatively robust to partitioning the data into these two samples. At -0.08, the estimated elasticity in the Philadelphia and Pittsburgh municipality regions is just slightly less than the elasticity in column 4 of Table 4. In the remaining regions (with data only available from 2004 onwards and excluding Pittsburgh and Philadelphia), the estimate is -0.15. These results appear to contrast with the results in Wadud et al. (2010a), which use U.S. survey data to find that drivers in urban areas in the United States are more responsive than those rural areas.

One explanation for this greater responsiveness may be that the sample only includes a later time period that has higher and more volatile gasoline prices. Lin and Prince (2013) provide evidence indicating that the price elasticity of gasoline demand is nearly twice as high during the 2007 to 2012 period (e.g., a period of high volatility) than the 2000 to 2006 period using aggregate monthly time series data for the United States. However, when we estimate the elasticity using the OBD 2000 vehicles (Philadelphia and Pittsburgh), but only for tests after 2005, we find only a slight change in the elasticity, with an estimated coefficient of -0.10 in the IV specification.

Another explanation is that the more rural areas have a different population and vehicle stock than the Philadelphia and Pittsburgh metropolitan areas. As shown in Table 2, the average age of vehicles is lower in the remaining regions than in metropolitan areas, perhaps suggesting that heterogeneity in the response by vehicle attributes may be an important part of the explanation.

## **5.2 Heterogeneity in responsiveness by quantile**

Quantile regressions are well-suited for examining the heterogeneity in the response across consumers. There are several approaches used in the literature to estimate quantile regression models with panel data. We adopt the approach of Canay (2011). The approach is quite simple.

First we perform our instrumental variable within-estimation with vehicle fixed effects ( $\theta_i$ ). We then estimate these individual fixed effects as the sample mean of the residual taken over each vehicle. Finally, we perform a standard quantile regression of  $\log(\widehat{VMT}_{it}) = \log(VMT_{it}) - \hat{\theta}_i$  on all of the covariates in (1), but with no vehicle fixed effects. Canay (2011) shows that this approach provides a consistent quantile fixed effects estimator as  $T \rightarrow \infty$  and  $n \rightarrow \infty$ .

By comparing the 0.25 quantile estimated elasticity to 0.5 (median) and 0.75 quantile elasticities in Table 5, we can see the breadth of responsiveness to gasoline price changes across the population. For computational reasons, we estimate the model on a randomly drawn subsample of 10% of the vehicles in the full dataset.

Table 5: Quantile regression results.

	(1)	(2)	(3)
	0.25 Quantile; IV	0.5 Quantile; IV	0.75 Quantile; IV
ln(gasoline price)	-0.020 (0.021)	-0.036*** (0.001)	-0.153*** (0.002)
ln(GDP)	6.232*** (0.376)	7.117*** (0.342)	4.483*** (0.375)
ln(unemployment)	0.835*** (0.053)	0.944*** (0.049)	0.692*** (0.053)
Age	-0.017*** (0.002)	-0.009*** (0.001)	0.001 (0.002)
age <sup>2</sup>	-0.002*** (7.48E-5)	-0.003*** (4.73E-5)	-0.005*** (6.80E-5)
vehicle fixed effects	X	X	X
year dummies	X	X	X
month-to-test dummies	X	X	X
driving period dummies	X	X	X
N	3,062,172	3,062,172	3,062,172

Notes: Dependent variable is the log average VMT over the driving period scaled to one year. The results are based on a randomly drawn subsample of 10% of the full dataset. All columns instrument for the log gasoline price with the log of the average size of the refinery disruptions over the driving period. Standard errors in parentheses. \*\*\* denotes significant at the 1% level. Vehicle fixed effects accounted for using the Canay (2011) quantile fixed effects estimator.

The results in Table 5 show clear heterogeneity in the response to gasoline prices. At the 0.25 quantile of the responsiveness, the elasticity is near-zero and statistically insignificant. The estimated elasticity from the median regression is statistically significant and closer to zero than our estimated elasticity in our mean regression in Table 3, at -0.04. On the other end, the 0.75 quantile elasticity is statistically significant and larger than our mean elasticity, at -0.15.

The heterogeneity we find in this setting is consistent with findings in Gillingham (2014), but the magnitude of the response is substantively less here. Such a difference may be due to differences in the sample: our sample contains most vehicles in Pennsylvania, rather than new vehicles in California. We now seek to explain some of this heterogeneity in responsiveness in two key vehicle attributes with policy relevance: fuel economy and age.

### 5.3 Heterogeneity in responsiveness by fuel economy

One might expect that consumers who drive less efficient vehicles would be more responsive to gasoline price changes since they face a greater burden at the pump. This would be particularly interesting for two reasons. First, if less efficient vehicles are more responsive, as fuel economy standards in the United States ramp up consumers may become even less responsive to changes in gasoline prices into the future. Second, if less efficient vehicles are more responsive, the emissions may be reduced by more than would be expected based on the mean elasticity, for less efficient vehicles have a higher emissions rate (Knittel and Sandler, 2013).

We examine heterogeneity based on vehicle fuel economy by estimating the model given in (1) on three partitions of the dataset. Our results are robust to the exact partitions we choose, so for ease of interpretability, Table 6 presents the results for three groups: vehicles with fuel economy less than 20 MPG (column 2), vehicles with fuel economy between 20 and 30 MPG (column 3), and vehicles with fuel economy greater than 30 MPG (column 4). Column 1 provides the estimation results using the full dataset for reference.

Table 6: Regression results on subsamples of the dataset based on fuel economy groups.

	(1)	(2)	(3)	(4)
	Full Dataset; IV	FE<20 MPG; IV	FE <sub>20-30</sub> MPG; IV	FE>30 MPG; IV
ln(gasoline price)	-0.099*** (0.008)	-0.192*** (0.013)	-0.008 (0.012)	-0.059 (0.067)
ln(GDP)	3.18***	3.92***	2.66***	3.94***

	(0.056)	(0.089)	(0.081)	(0.525)
ln(unemployment)	0.328***	0.380***	0.313***	0.426***
	(0.004)	(0.007)	(0.006)	(0.044)
age	-0.091***	-0.0926***	-0.0957***	-0.112***
	(7.38E-4)	(0.001)	(0.001)	(0.008)
age <sup>2</sup>	-1.79E-3***	-2.61E-3***	-7.35E-4***	-9.81E-4***
	(1.52E-5)	(2.16E-5)	(2.12E-5)	(1.54E-5)
vehicle fixed effects	X	X	X	X
year dummies	X	X	X	X
month-of-year dummies	X	X	X	X
driving period dummies	X	X	X	X
R-squared	0.05	0.06	0.03	0.03
N	30,621,721	14,773,863	15,555,437	324,310

Notes: Dependent variable is the log average VMT over the driving period scaled to one year. The log gasoline price is instrumented using the log refinery disruptions in all columns. Column 1 is estimated on the full dataset (identical to column 3 in Table 3), 2 on vehicles with a fuel economy below 20 MPG, 3 on vehicles with a fuel economy 20-30 MPG, and 4 on vehicles with fuel economy greater than 30 MPG. R-squared values are the within R-squared. Robust standard errors in parentheses, clustered at the vehicle level. \*\*\* denotes significant at the 1% level.

The results from Table 6 show that drivers of vehicles in the lowest fuel economy bracket are the most responsive, with an elasticity of -0.19. Interestingly, the coefficient on the gasoline price is close to zero and statistically insignificant in columns 3 and 4, suggesting that drivers of vehicles of higher fuel economy are barely responsive to changes in gasoline prices. These results provide strong evidence that the lowest fuel economy vehicles are the force behind the responsiveness to gasoline prices.

There are several possible explanations for this result. One is that gasoline prices are simply more salient to consumers of lower fuel economy vehicles, due to their higher fuel bill. Another is the drivers of lower fuel economy vehicles are less wealthy and thus face a tighter budget constraint. Still another is that there may be within-household vehicle switching occurring, consistent with previous results in the literature, such as in Knittel and Sandler (2013). All these

effects could also be occurring concurrently, and we cannot disentangle them with our data. However, future work to understand these mechanisms could be valuable.

#### **5.4 Heterogeneity in responsiveness by age of vehicles**

Understanding how the responsiveness to changing gasoline prices differs by the age and cohort of vehicles is useful for both forecasting driving by the fleet, as well as determining the changes in local air pollution from a policy. Specifically, many older vehicles are more polluting than new vehicles (although within the mandated guidelines from emissions inspections). If older vehicles are more responsive to changes in gasoline prices, this may imply that meeting local air pollution targets will be disproportionately easier when gasoline prices increase.<sup>12</sup>

Table 7 provides similar results to previous tables, only using subsamples of the dataset based on the vehicle age. Again for ease of interpretation we choose three subsamples: vehicles with age less than 3 years, age between 3 and 7 years, and age greater than 7 years. The logic for these cut-offs is that vehicles less than three years old are very new and often leased, vehicles between 3 and 7 years make up the bulk of VMT of owned vehicles, while vehicles older than 7 years are driven less. Columns 2, 3, and 4 in Table 7 present the results from each of these subsets of the data respectively. Column 1 again shows the results estimated on the full dataset for reference.

---

<sup>12</sup> Note that the average fuel economy of the fleet has not changed significantly over the 2000-2010 time period; it has hovered around 21 mpg in all of those years.

Table 7: Regression results on subsamples of the dataset based on vehicle age.

	(1)	(2)	(3)	(4)
	Full Dataset; IV	Age<3; IV	Age3-7; IV	Age>7; IV
ln(gasoline price)	-0.099*** (0.008)	-0.083* (0.039)	-0.411*** (0.013)	-0.007 (0.013)
ln(GDP)	3.18*** (0.056)	2.94*** (0.342)	8.07*** (0.086)	0.357*** (0.082)
ln(unemployment)	0.328*** (0.004)	0.481*** (0.032)	0.654*** (0.007)	0.026*** (0.007)
age	-0.091*** (7.38E-4)	0.391*** (0.106)	-0.145*** (0.002)	-0.059*** (0.001)
age <sup>2</sup>	-1.79E-3*** (1.52E-5)	-0.150*** (0.005)	-0.003*** (1.33E-4)	-8.22E-4*** (3.22E-5)
vehicle fixed effects	X	X	X	X
year dummies	X	X	X	X
month-of-year dummies	X	X	X	X
driving period dummies	X	X	X	X
R-squared	0.05	0.02	0.02	0.04
N	30,621,721	1,741,169	10,186,019	15,409,706

Notes: Dependent variable is the log average VMT over the driving period scaled to one year. The log gasoline price is instrumented using the log refinery disruptions in all columns. Column 1 is estimated on the full dataset (identical to column 4 in Table 3), 2 on vehicles with a fuel economy below 20 MPG, 3 on vehicles with a fuel economy 20-30 MPG, and 4 on vehicles with fuel economy greater than 30 MPG. R-squared values are the within R-squared.

Robust standard errors in parentheses, clustered at the vehicle level. \*\*\* denotes significant at the 1% level, and \* at the 10% level.

Our results indicate that drivers of vehicles that are 3 to 7 years old are the most responsive to changes in gasoline prices, with an elasticity of -0.41. Vehicles in these age brackets are driven the most, so perhaps there is the most room to make changes. In contrast, the newest vehicles have an elasticity of -0.08, which is not quite as statistically significant. Drivers of the newest vehicles may be wealthier, which may influence the elasticity. The oldest vehicles have an elasticity very close to zero and statistically insignificant. There are a variety of possible explanations for this result. For example, one possibility is that drivers of older vehicles may be more likely to be in rural areas, so they must drive more for work. Another is the older vehicles are driven so much less in general that there are fewer “flexible” miles to reduce.

## **5.5 Price per mile results**

Our previous results are based on specification (1), which models VMT as a function of the gasoline price. However, it is perhaps natural to think of the price of VMT to be the price per mile of driving, which is the price of gasoline divided by the fuel economy. In fact, examining the elasticity of VMT with respect to the price per mile of driving is closer to what we are ideally looking for in examining the rebound effect from a policy that increases fuel economy. If consumers respond in the same way to changes in gasoline prices and changes in fuel economy—and there are no close substitutes to driving that use gasoline—then one would expect the elasticity of VMT with respect to gasoline prices, the price per mile of driving, and fuel economy to be the same (Chan and Gillingham, 2015). For this reason alone, it is instructive to examine the elasticity of VMT with respect to the price per mile of driving.<sup>13</sup>

We thus estimate (1) using the price per mile of driving rather than the gasoline price. Of course, the variation identifying the coefficient on the price per mile of driving is identical to the variation in our primary specifications, so we should expect to see very similar results. Table 8 shows that the results are indeed nearly the same as those in Table 3 and Table 4.

---

<sup>13</sup> We cannot examine the elasticity with respect to fuel economy while including vehicle fixed effects, which we deem essential to control for selection into vehicles based on fuel economy and location-specific unobservables.

Table 8: Regression results examining the average price per mile (2010\$/mi).

	(1)	(2)	(3)
	Full Dataset; IV	OBD 2000; IV	OBD 2004; IV
ln(price per mile driven)	-0.097*** (0.009)	-0.076*** (0.015)	-0.150*** (0.02)
ln(GDP)	3.28*** (0.060)	3.33*** (0.082)	4.31*** (0.240)
ln(unemployment)	0.345*** (0.005)	0.398*** (0.006)	0.429*** (0.014)
Age	-0.0920*** (7.72E-4)	-0.0973*** (9.25E-4)	-0.0891*** (0.002)
age <sup>2</sup>	-1.79E-3*** (1.52E-5)	-1.767E-3*** (1.84E-5)	-2.42E-3*** (3.98E-5)
vehicle fixed effects	X	X	X
year dummies	X	X	X
month-of-year dummies	X	X	X
driving period dummies	X	X	X
R-squared	0.04	0.05	0.05
N	30,653,610	19,938,566	3,163,074

Notes: Dependent variable is the log average VMT over the driving period scaled to one year. The log price per mile is instrumented using the log refinery disruptions. Column 1 is estimated on the full dataset (identical to column 3 in Table 3), 2 on the subsample that first required OBD inspections in 2000 (i.e., Philadelphia and Pittsburgh metro areas), and 3 on the subsample that first required OBD inspections in 2004. R-squared values are the within R-squared. Robust standard errors in parentheses, clustered at the vehicle level. \*\*\* denotes significant at the 1% level.

Our results in Table 8 are consistent with Frondel and Vance (2013), who also finds little difference regardless of the choice of the key independent variable used to identify an elasticity. However, it differs from Linn (2013), which uses different sources of variation to identify the different elasticities and finds a greater responsiveness to fuel economy than to gasoline prices. We interpret our result as providing further evidence suggestive of a moderate short-run rebound effect in Pennsylvania, consistent with the results for the United States in Hymel and Small (2014).

## **6 Conclusions**

The price elasticity of VMT demand has long been an object of interest for both economists and policymakers. Besides providing guidance for understanding the effects of changing gasoline prices and taxes, it also provides insight into the rebound effect in personal transportation. With policies to improve fuel economy remaining at the core of energy and transportation policy in many countries, understanding the potential magnitude of the rebound effect is essential for analyses of these policies, and the gasoline price elasticity is commonly used to provide guidance on the rebound effect.

This paper sheds new light on the gasoline price elasticity of VMT demand by examining the heterogeneity in responsiveness, and some of the most interesting factors underlying this heterogeneity. We use rich annual vehicle inspection data of over 30 million observations from the state of Pennsylvania, covering the period of 2000 to 2010. Using the revealed odometer readings from the tests, we find that the vehicle-level response to gasoline price changes is highly inelastic. In fact, our primary results suggest a one-year gasoline price elasticity of VMT demand of -0.10. This result is in line with much of the recent previous literature that shows a highly inelastic response to gasoline prices in the United States. We also find interesting heterogeneity in this mean elasticity, with quantile regression results indicating that a relatively small percentage of vehicles drive the elasticity.

To understand this heterogeneity more deeply, we explore how the responsiveness varies by fuel economy and age of vehicles. First, our results indicate substantial heterogeneity across fuel

economy in the fleet. The highest fuel economy vehicles display an entirely inelastic response and nearly all of the response is from lower fuel economy vehicles. There are several possibilities that could explain this finding. One is that when gasoline prices rise, consumers switch from lower fuel economy vehicles to higher fuel economy vehicles. Another is that consumers with lower fuel economy vehicles may be expected to respond more due to the greater salience of gasoline costs or have a greater income effect.

Our results indicate an intriguing pattern of responsiveness across vehicle age. We find that the newest and oldest vehicles are the least responsive to gasoline price changes. It appears that consumers driving vehicles 3 to 7 years old are the most sensitive to gasoline prices. While we cannot definitely determine the mechanism underlying this result, it may be that drivers of new vehicles (many of which are leased) are highly inelastic, while drivers of the oldest vehicles have less flexibility in travel. The findings of heterogeneity based on both fuel economy and age have important implications for forecasting the usage of the vehicle fleet as gasoline prices change, and are also useful for forecasting local air pollution due to the greater local air pollutant emissions per mile from lower fuel economy vehicles and older vehicles.

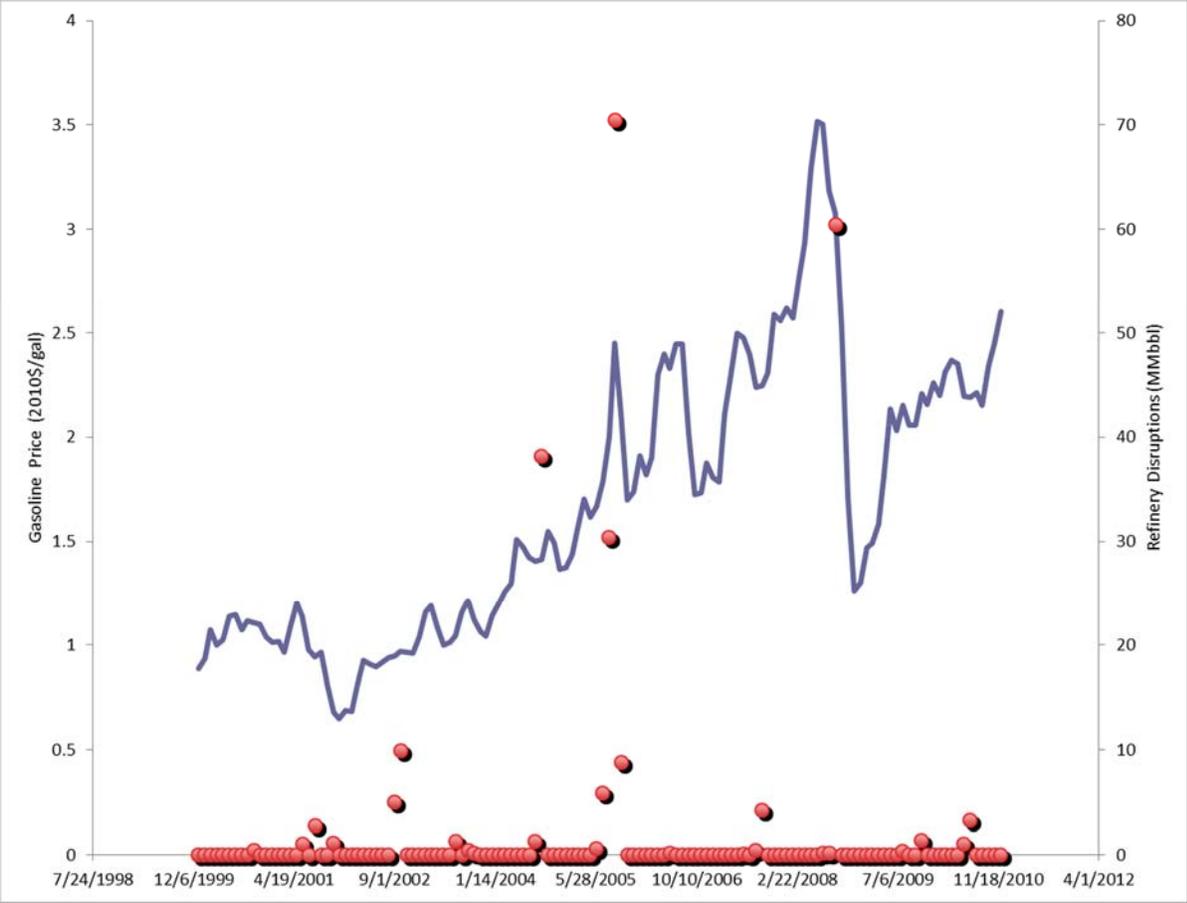
The result of an inelastic response, but considerable heterogeneity in this response, also provides some guidance for the short-run direct rebound effect from fuel economy standards and other policies to improve fuel economy, such as feebates. Following the common assumption that the consumer response to fuel prices is the same as the response to changes in fuel economy, our results suggest that the short-run direct rebound effect may be relatively small in Pennsylvania. Furthermore, our results also suggest that it may be larger for some vehicles than others. Such heterogeneity in the rebound effect may be very important for policies that asymmetrically improve fuel economy across the fleet. For example, if the direct rebound effect is close to zero for higher fuel economy vehicles and fuel economy standards disproportionately improve the fuel economy of higher fuel vehicles, then the average rebound effect across the fleet can be expected to be smaller. Exploring further details of the consumer responsiveness to improvements in the fuel efficiency of the light duty fleet promises to be a valuable area for future research.

## Online Appendix

This online appendix presents the first stage regression results of our primary IV results and a graph showing the variation in gasoline prices and refinery disruptions.

	(1)
	Primary
ln(refinery disruptions)	0.011*** (6.46E-6)
ln(GDP)	5.331*** (0.002)
ln(unemployment)	-0.130*** (0.002)
age	-0.011*** (4.39E-5)
age <sup>2</sup>	-5.92E-5*** (4.06E-7)
vehicle fixed effects	X
year dummies	X
month-of-year dummies	X
driving period dummies	X
F-statistic	2.8x10 <sup>6</sup>
R-squared	0.97
N	30,621,721

Notes: Dependent variable is the log average gasoline price over a driving period. Robust standard errors clustered on vehicle in parentheses. \*\*\* denotes significant at the 1% level.



## References

- Azevedo IL. 2014. Consumer End-Use Energy Efficiency and Rebound Effects. *Annual Review of Environment and Resources*, 39, 393-418.
- BEA. 2014. U.S. Department of Commerce, Bureau of Economic Analysis, GDP. Available at: <http://www.bea.gov/national/index.htm>.
- BLS. 2014. Bureau of Labor Statistics: Unemployment Rates for States. Available at: <http://data.bls.gov/timeseries/LNS14000000>.
- Blundell R, Horowitz J, Parey M. 2012. Measuring the Price Responsiveness of Gasoline Demand: Economic Shape Restrictions and Nonparametric Demand Estimation. *Quantitative Economics*, 3, 29-51.
- Canay I. 2011. A Simple Approach to Quantile Regression for Panel Data. *The Econometrics Journal*, 14, 368-386.
- Chan N, Gillingham K. 2015. The Microeconomic Theory of the Rebound Effect and its Welfare Implications. *Journal of the Association of Environmental and Resource Economists*, 2, 133-159.
- Copeland A, Dunn W, Hall G. 2011. Inventories and the Automobile Market. *RAND Journal of Economics*, 42, 121-149.
- Dahl C, Sterner T. 1991. Analysing Gasoline Demand Elasticities: A Survey. *Energy Economics*, 13, 203-210.
- EIA. 2011. US Energy Information Administration. Independent Statistics & Analysis. Short-Term Energy Outlook Supplement: 2011 Outlook for Hurricane-Related Production Outages in the Gulf of Mexico.
- EIA. 2014. U.S. Energy Information Administration, Historical Crude Oil Prices. Available at: [http://www.eia.gov/dnav/pet/pet\\_pri\\_spt\\_s1\\_m.htm](http://www.eia.gov/dnav/pet/pet_pri_spt_s1_m.htm).
- EPA. 2014. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2012. EPA 430-R-14-003. Available at: <http://www.epa.gov/climatechange/Downloads/ghgemissions/US-GHG-Inventory-2014-Main-Text.pdf>.
- Frondel M, Ritter N, Vance C. 2012. Heterogeneity in the Rebound Effect: Further Evidence for Germany. *Energy Economics*, 34, 461-467.
- Frondel M, Vance C. 2013. Re-Identifying the Rebound: What about Asymmetry? *Energy Journal*, 34, 43-54.
- Gately D, Huntington H. 2002. The Asymmetric Effects of Changes in Price and Income on Energy and Oil Demand. *Energy Journal*, 23, 19-55.
- Gillingham K, 2011. The Consumer Response to Gasoline Price Changes: Empirical Evidence and Policy Implications. Stanford University Ph.D. Dissertation.
- Gillingham K. 2013. Selection on Anticipated Driving and the Consumer Response to Changing Gasoline Prices. Yale University Working Paper.
- Gillingham K. 2014. Identifying the Elasticity of Driving: Evidence from a Gasoline Price Shock. *Regional Science & Urban Economics*, 47, 13-24.

- Gillingham K, Kotchen M, Rapson D, Wagner G. 2013. The Rebound Effect is Over-played. *Nature*, 493, 475-476.
- Gillingham K, Munk-Nielsen A. 2015. The Tail Wagging the Dog: Fuel Prices and Driving Decisions in Denmark. Yale University Working Paper.
- Gillingham K, Rapson D, Wagner G. 2016. The Rebound Effect and Energy Efficiency Policy. *Review of Environmental Economics & Policy*, forthcoming.
- Goodwin P, Dargay J, Hanly M. 2004. Elasticities of Road Traffic and Fuel Consumption with Respect to Price and Income: A Review. *Transport Reviews*, 24, 275-292.
- Graham D, Glaister S. 2004. Road Traffic Demand Elasticity Estimates: A Review. *Transport Reviews*, 24, 261-274.
- Greening LA, Greene DL, Difiglio C. 2000. Energy efficiency and consumption—the rebound effect—a survey. *Energy Policy*, 28, 389-401.
- Hausman J, Newey W. 1995. Nonparametric Estimation of Exact Consumers Surplus and Deadweight Loss. *Econometrica*, 63, 1445-1476.
- Hughes J, Knittel C, Sperling D. 2008. Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand. *Energy Journal*, 29, 93-114.
- Hymel KM, Small KA. 2014. The Rebound Effect for Automobile Travel: Asymmetric Response to Price Changes and Novel Features of the 2000s. *Energy Economics*, 49, 93-103.
- Jenn A, Azevedo IL, Fischbeck P. 2015. How Will We Fund Our Roads? A Case of Decreasing Revenue From Electric Vehicles. *Transportation Research A*, 74, 136-147.
- Kennan J. 1989. Simultaneous Equations Bias in Disaggregated Econometric Models. *Review of Economic Studies*, 56, 151-156.
- Knittel C, Sandler R. 2013. The Welfare Impact of Indirect Pigouvian Taxation: Evidence from Transportation. National Bureau of Economic Research Working Paper No. 18849.
- Koenker R, Bassett G. 1978. Regression Quantiles. *Econometrica*, 46, 33-50.
- Li S, Linn J, Muehlegger E. 2014. Gasoline Taxes and Consumer Behavior. *American Economic Journal: Economic Policy*, 6, 302-342.
- Lin C-YC, Prince L. 2013. Gasoline Price Volatility and the Elasticity of Demand for Gasoline. *Energy Economics*, 38, 111-117.
- Linn J. 2013. The Rebound Effect for Passenger Vehicles. RFF Working Paper.
- Liu W. 2014. Modeling Gasoline Demand in the United States: A Flexible Semiparametric Approach. *Energy Economics*, 45, 244-253.
- PennDOT. 2010. Pennsylvania DOT Summary of Vehicle Registrations, <http://www.dmv.state.pa.us/pdotforms/stats/ReportofRegistration2010.pdf>.
- Schmalensee R, Stoker T. 1999. Household Gasoline Demand in the United States. *Econometrica*, 67, 645-662.
- Small KA, van Dender K. 2007a. Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect. *Energy Journal*, 28, 25-51.

- Small KA, Van Dender K. 2007b. Long Run Trends in Transport Demand, Fuel Price Elasticities and Implications of the Oil Outlook for Transport Policy. OECD/ITF Joint Transport Research Centre Discussion Paper No. 2007-16.
- Sorrell S. 2007. The Rebound Effect: An Assessment of the Evidence for Economy-wide Energy Savings from Improved Energy Efficiency. UK Energy Research Center Report.
- Thomas BA, Azevedo IL. 2013. Estimating direct and indirect rebound effects for US households with input–output analysis Part 1: Theoretical framework. *Ecological Economics*, 86, 199-210.
- Thomas BA, Azevedo IL. 2014. Should Policy-makers Allocate Funding to Vehicle Electrification or End-use Energy Efficiency as a Strategy for Climate Change Mitigation and Energy Reduction? Rethinking Electric Utilities Efficiency Programs. *Energy Policy*, 67, 28-36.
- Wadud Z, Graham D, Noland R. 2010a. Gasoline Demand with Heterogeneity in Household Responses. *Energy Journal*, 31, 47-74.
- Wadud Z, Noland R, Graham D. 2010b. A Semiparametric Model of Household Gasoline Demand. *Energy Economics*, 32, 93-101.
- Yatchew A, No JA. 2001. Household Gasoline Demand in Canada. *Econometrica*, 69, 1697-1709.