Federal Reserve Bank of New York
Staff Reports

Do Long-Haul Truckers Undervalue Future Fuel Savings?

Jacob Adenbaum
Adam Copeland
John Stevens

Staff Report No. 756
January 2016
Revised July 2016

FRBNY Staff REPORTS

Do Long-Haul Truckers Undervalue Future Fuel Savings?
Jacob Adenbaum, Adam Copeland, and John Stevens
Federal Reserve Bank of New York Staff Reports, no. 756
January 2016; revised July 2016
JEL classification: D22, L51, L92

Abstract

The U.S. federal government enacted fuel efficiency standards for medium and heavy trucks for the first time in September 2011. Rationales for using this policy tool typically depend upon frictions existing in the marketplace or consumers being myopic, such that vehicle purchasers undervalue the future fuel savings from increased fuel efficiency. We measure by how much long-haul truck owners undervalue future fuel savings by employing recent advances to the classic hedonic approach to estimate the distribution of willingness-to-pay for fuel efficiency. We find significant heterogeneity in truck owners' willingness-to-pay for fuel efficiency, with the elasticity of fuel efficiency to price ranging from 0.51 at the 10th percentile to 1.33 at the 90th percentile, and an average of 0.91. Combining these results with estimates of future fuel savings from increases in fuel efficiency, we find that long-haul truck owners' willingness-to-pay for a 1 percent increase in fuel efficiency is, on average, just 29.5 percent of the expected future fuel savings. These results suggest that introducing fuel efficiency standards for heavy trucks might be an effective policy tool to raise medium and heavy trucks' fuel economy.

Key words: fuel efficiency standards, durable goods, discrete-choice demand estimation

---

Adenbaum, Copeland: Federal Reserve Bank of New York (e-mail: jacob.adenbaum@ny.frb.org, adam.copeland@ny.frb.org). Stevens: Federal Reserve Board (e-mail: john.j.stevens@frb.gov). Address correspondence to Adam Copeland. The authors thank Ed McCarty and John Miller for assistance with the data, Eugene Chung and Sarit Weisburd for research assistance, and Kenneth Gillingham and Jason Saving for comments.

1 Introduction

In September 2011, the U.S. federal government set fuel efficiency standards for medium and heavy trucks for the first time, and in June 2015, the National Highway Transportation and Safety Administration (NHTSA) and the Environmental Protection Agency (EPA) proposed a second round of standards.[1] The motivation for these policy interventions is to reduce air pollution and other negative externalities associated with the use of medium and heavy trucks, similar to the arguments made to set fuel efficiency standards for passenger cars.

The effectiveness of using fuel efficiency standards to reduce negative externalities is an open question (Parry et al., 2007). In particular, some economists argue that fuel taxes are more effective at reducing air pollution (e.g., see West and Williams III (2005) and references therein). As a result, the motivation for mandating fuel efficiency standards often relies upon the existence of marketplace frictions or consumer myopia, both of which can lead to vehicle purchasers undervaluing future fuel savings and vehicle producers under-investing in technologies that improve fuel efficiency.

In presenting the new regulations, the NHTSA and the EPA made the case that the current adoption of fuel savings technologies is inefficiently low by arguing that there are technologies available for which the cost of adoption is lower than the expected discounted value of future fuel savings (NHTSA-EPA, 2015). Further, for both the 2011 standards and the proposed 2015 standards, the agencies emphasized that the anticipated fuel savings from meeting these standards is (much) greater than the expected costs.[2]

The NHTSA and the EPA then motivate why there might be inefficienct adoption of fuel savings technologies by listing a number of potential reasons why medium and heavy truck owners might undervalue future fuel savings. For example, imperfect information among truck owners about the effectiveness of potential fuel efficiency technologies could make owners unwilling to purchase these technologies. Alternatively, truck owners may be myopic with respect to future fuel costs and therefore unwilling to pay upfront for a technology that delivers fuel savings in the future. Finally, agency problems may exist whereby the firms that purchase the trucks do not always pay the fuel costs, and drivers are not rewarded for operating in a fuel-efficient manner (e.g., see Vernon and

[1] For details on the proposed 2015 rulemaking, see `http://www.epa.gov/oms/climate/regs-heavy-duty.htm`.

[2] For example, the NHTSA and the EPA argue that the typical buyer of a new long-haul truck in 2027 (when the second set of proposed standards are fully phased in) will recoup the extra costs of the fuel efficiency technologies in under 2 years (see the NHTSA-EPA June 2015 press release at `http://www.epa.gov/otaq/climate/documents/420f15900.pdf`).

Meier (2012)).

As a first step toward understanding whether the new fuel efficiency standards for medium and heavy trucks might be an effective policy instrument, we look for evidence that truck owners undervalue future fuel savings from increased fuel efficiency. Rather than consider all medium and heavy trucks, we focus on class 8 long-haul trucks, which are the large truck-tractors used on highways that specialize in hauling cargo long distances. We do so because fuel efficiency is a main cost for owners of these vehicles and so should be a characteristic to which they are attuned.[3] Consequently, a long-haul truck's fuel efficiency should be reflected in its price. Although these trucks are not representative of medium and heavy trucks, long-haul trucks, which are a major part of the national transportation system for freight, disproportionately account for fuel consumption in the medium and heavy truck segment because of their high annual miles traveled and heavy payloads.[4] Thus, the overall effectiveness of fuel economy standards depends, in large part, on the long-haul truck market. In addition, long-haul trucks are clearly defined in the data and contain a relatively small number of producers and products, which facilitates our market analysis. Lastly, there exist detailed data on the stock and value of used long-haul trucks that enables us to perform our empirical analysis.

We use data from two sources. The Census Bureau's Vehicle Inventory and Use Survey (VIUS) provides information on the stock and characteristics of all medium and heavy trucks registered in the United States. These data provide an extraordinary amount of information on trucks, with survey respondents, i.e. truck owners, answering questions on the physical and operational characteristics of their truck.[5] We use the reported truck characteristics both to identify class 8 long-haul trucks in the VIUS and to merge information on used truck prices from the *Truck Blue Book*. The end result of this merger is a data set on the equilibrium price and quantity for a representative sample of long-haul trucks registered in the United States. We have merged the VIUS and *Truck Blue Book* prices for 1992 and 1997. Our results focus on 1992, although we show the results are robust to using the 1997 data.

[3] For example, Torrey IV and Murray (2014) (page 6) report that fuel costs are "one of the top two cost centers for motor carriers."

[4] The NHTSA and the EPA report that class 7 and class 8 trucks account for 65 percent of fuel consumption in the heavy-duty sector (NHTSA-EPA (2011), page 11).

[5] These data have been used to study a variety of empirical questions, e.g., productivity (Hubbard, 2003), technology adoption and governance (Baker and Hubbard, 2004), and merger analysis which takes into account the entry and exit of products (Wollmann, 2014).

We use these data to estimate truck owners' willingness-to-pay for truck characteristics, employing recent advances to the hedonic approach laid out in Bajari and Benkard (2005). Their approach allows the econometrician to recover willingness-to-pay in an environment with imperfect competition and a product characteristic that is unobserved (by the econometrician); both of these features are present in the heavy truck market. We employ a local quadratic method to infer truck owners' willingness-to-pay for four continuous characteristics—miles-per-gallon (MPG), lifetime miles, engine size, and the truck's weight when empty—while controlling for a number of fixed effects. This method allows us to recover nonparametrically the distribution of willingness-to-pay for each continuous characteristic.

Our estimates of these distributions have the expected signs in that almost all truck owners negatively value lifetime miles and positively value engine size, empty weight, and MPG. In particular, we find that the mean elasticity of MPG to price is 0.91, or that on average a truck owner is willing to pay 0.91 percent of his truck's price for a 1 percent increase in MPG, holding all else equal. The 10th and 90th percentiles of this distribution are 0.51 and 1.33, demonstrating that there is a wide range of tastes for fuel efficiency among truck owners.

To determine whether long-haul truck owners undervalue expected savings from fuel efficiency, we need to compute the discounted future savings from a 1 percent increase in MPG. These savings vary across trucks, depending upon their characteristics (such as current age and fuel efficiency). We then compare these estimated savings with our estimates of willingness-to-pay by looking at the ratio of WTP for a 1 percent increase in fuel efficiency over the expected future fuel savings from a 1 percent increase in fuel efficiency, for each truck owner. We find that the mean of this ratio is 29.5 percent, which means that trucks owners are willing to pay only 29.5 cents today for a dollar in expected fuel savings (in net present value terms). The average truck owner then, undervalues the discounted future savings from increased fuel efficiency. There is substantial variation in this measure across truck owners; the distribution of the ratio of the willingness-to-pay over future fuel savings ranges from 8.8 percent at the 10th percentile to 54.5 percent at the 90th percentile. The main result of this paper, then, is that almost all long-haul truck owners undervalue expected lifetime fuel savings, which in turn suggests that imposing fuel efficiency standards could be an effective policy tool to raise the fuel economy of long-haul trucks.

A second set of results evaluates the willingness of new truck owners to adopt a suite of technologies that would dramatically improve fuel efficiency. The NHTSA and the EPA argue that technologies exist today that make it feasible for heavy duty engine and truck manufacturers to produce trucks which meet the fuel efficiency targets mandated in the 2011 final ruling. A concrete example is provided by the SuperTruck program, a government-sponsored research project whose goal is to demonstrate that a 50 percent improvement in fuel efficiency for class 8 long-haul trucks is feasible. The final report of the SuperTruck program provides tables on both the expected fuel efficiency gains from introducing two different sets of innovations as well as the incremental costs of both packages of innovations.[6] The first set of innovations is predicted to increase MPG of a current truck-tractor model by 65.3 percent, at the cost of raising the price of the truck by 26.6 percent. The second set of innovations provides better fuel efficiency, increasing MPG by 69.8 percent, but at a much higher cost of a 51.0 percent increase over the current truck price. Using this information and our estimates of willingness to pay, we can calculate the fraction of long-haul truckers purchasing (almost) new trucks for which the willingness-to-pay for fuel efficiency is higher than the incremental costs. We find that 93.6 percent of these long-haul truck owners would be willing to adopt the first set of innovations, and 80.9 percent of owners would be willing to adopt the second set. These results are encouraging then, in that the model predicts that a strong majority of new long haul truck owners, despite their undervaluation of future fuel savings, would be willing to bear the costs of adopting the SuperTruck set of fuel efficiency innovations.

From a policy perspective, a caveat of our results is that our analysis is based on prices and quantities in 1992 whereas the new fuel efficiency standards are being introduced in 2015, more than twenty years later. Our results, however, are based on the structural parameters of the truck owners' problem, and it is not unusual to assume that deep parameters do not change over periods of this length. Furthermore, in the robustness section we demonstrate that our results also hold using 1997 data.[7] Nevertheless, if it were the case that there have been substantial changes in the long-haul trucking industry, our results should be viewed with the appropriate amount of skepticism.

Our contribution to the literature on evaluating the effectiveness of fuel efficiency standards is twofold. First, we focus on class 8 long-haul trucks whereas previous work analyzed automobiles and light trucks, as U.S. fuel efficiency standards applied to only this subset of vehicles prior to

[6]The SuperTruck final report can be found on the Department of Transportation's website, at `http://www.transportation.anl.gov/pdfs/TA/903.PDF`.

[7]The VIUS was last conducted in 2002. However, as explained in section 3.1, we merge the VIUS data with the *Truck Blue Book* prices using a 2-digit VIN. The Census Bureau provided us with a special tabulation of the 2-digit VIN for each truck in the 1992 and 1997 VIUS, but declined our follow-up request for the same information from the 2002 VIUS.

2011.[8] However, the recent expansion of fuel efficiency standards to medium and heavy trucks raises similar questions about the effectiveness of this policy. To the best of our knowledge, this paper is the first to formally test whether long-haul truckers undervalue the expected future discounted gains from increased fuel efficiency. Our main result—that these truckers considerably undervalue future fuel savings—suggests that fuel efficiency standards could be effective tools when applied to long-haul trucks.[9]

The second contribution of our paper is our econometric approach, which has not yet been used in this literature.[10] We use a technique that has the advantage of allowing us to nonparametrically recover the distribution of willingness-to-pay in the population. Specifically, we use a local quadratic method to estimate long-haul truck owners' willingness to pay for fuel efficiency (and other characteristics). We then aggregate these estimates across truck owners to arrive at an estimate of the distribution in the population. Allowing for heterogeniety is recognized as an important feature when estimating demand for products generally (e.g., see the beginning of section 1 in Ackerberg et al. (2007)). Further, for the specific case of estimating willingness to pay for fuel efficiency, Bento et al. (2012) and Grigolon et al. (2014) argue that it is crucial to account for heterogeneity among consumers.

In the literature focused on fuel efficiency, there are empirical papers that allow for heterogeneity in consumer tastes; for example, Goldberg (1998) and Grigolon et al. (2014) estimate discrete-choice models with random coefficients. These approaches, however, require both an assumption about the distribution of willingness-to-pay and a panel data set.[11] With our approach, we can recover the willingness-to-pay distribution nonparametrically, and we can do so using only a cross-section of data.

As discussed in more detail in section 2, a main identifying assumption we make is that a truck's unobserved characteristic is independent of the truck's other (observed) characteristics. This assumption is slightly stronger than the mean independence assumption that is commonly used in

[8]Parry et al. (2007) and Helfand and Wolverton (2011) provide recent reviews of this literature.

[9]Leard et al. (2015) also consider how effective are the introduction of medium and heavy truck emissions standards. But they focus on estimating the rebound effect for these trucks and interpreting how their result changes the estimated long-run benefits from introducing the emissions standards.

[10]In a search of the empirical literature, we have found only three instances of published articles that implement this technique: Bajari and Kahn (2005), Koster et al. (2014), and Thomassen (2014).

[11]Researchers typically assume that willingness-to-pay, or tastes for characteristics, are drawn from a Gaussian distribution. These models can also be estimated with data on several markets (e.g. cities) at one point in time.

the empirical industrial organization literature.[12] Nevertheless, we argue later in the paper that this assumption is reasonable when considering the market for long-haul trucks. Further, we note here that one of many differences between our approach and some of the reduced form papers in this literature, is that the latter have been able to estimate consumers' willingness to pay under weaker identification assumptions, albeit with greater data demands. (See, in particular, Allcott and Wozny (2014) and Busse et al. (2013).)

In the next section, we introduce our model, and in section 3 we describe the data. Section 4 presents our empirical approach and the results on truckers' willingness-to-pay for fuel efficiency and other characteristics. In section 5, we compare the willingness-to-pay estimates for fuel efficiency against our measures of the expected discounted lifetime savings from increased fuel efficiency. This comparison reveals by how much truckers' undervalue expected discounted fuel savings. Section 6 concludes.

2 Model

In this section, we present our model of demand for purchasing long-haul trucks.[13] A long-haul truck is described by two types of attributes: physical attributes observed by both truck owners and the econometrician and a scalar characteristic which is observed only by truck purchasers. The physical characteristics used in our analysis include four continuous characteristics and a set of dummy variables that account for the truck manufacturer as well as the truck cab design. The continuous characteristics are MPG, lifetime miles, engine size (measured by cubic inch displacement), and empty weight (measured in pounds). By its nature, the unobserved characteristic is difficult to describe, but likely reflects a hard to measure attribute such as quality.

Following the notation of Bajari and Benkard (2005), let $j \in J$ index the trucks and $i \in I$ index truck owners. Suppose $\mathbf{x}_j$ denotes a $1 \times K$ vector of physical characteristics, p_j is the price of the truck, and ξ_j is the unobserved characteristic. A truck owner i maximizes utility by selecting a

[12]For example, Berry et al. (1995) and most models utilizing their estimation method assume that the unobserved characteristic is mean independent of the observed characteristics. See Bajari and Benkard (2005) for a detailed comparison of these two econometric approaches.

[13]Because we are using a technique from the random utility model literature, we talk about the truck owner's utility. However, the truck owner could equivalently be choosing among the many alternatives with the goal of maximizing profits.

product j as well as a composite good $c \in \mathbb{R}_{+}$. Truck owners have income y_i. Normalizing the price of c to 1, the truck owner's maximization problem is

$$\max_{(j,c)} u_i(\mathbf{x}_j, \xi_j, c)$$
$$\text{subject to } \; p_j + c \leq y_i.$$

Under fairly general conditions on u_i, Bajari and Benkard (2005) prove there is an equilibrium price function $\tilde{p}(\mathbf{x}, \xi_j)$ which maps the set of product characteristics to prices and satisfies $p_j = \tilde{p}(\mathbf{x}_j, \xi_j)$ for all $j = 1, \ldots, J$ for a specific market and point in time.[14] They also show that the unobserved characteristics can be be identified using a single cross-section: If $(p, \mathbf{x}, \xi)$ are distributed jointly with cumulative distribution function $F(p, \mathbf{x}, \xi)$, then the unobserved characteristics ξ_j are equal to the conditional cumulative distribution function of the prices,

$$\xi_j = F_{p|\mathbf{x}=\mathbf{x_j}}(p_j). \tag{1}$$

To identify the price function and the unobserved characteristics, we assume ξ is independent of $\mathbf{x}$. An interpretation of this assumption is that the location of trucks in the observed characteristic space is exogenous to ξ, or at least determined prior to the revelation of consumers' willingness-to-pay for ξ. This assumption is reasonable because truck and engine re-designs, which involve substantial R&D, are often done every several years whereas ξ can vary on a more frequent basis.[15]

Because we observe a single cross-section of truck owners, identification requires us to specify u_i; we make the fairly standard assumption that

$$u_{ij} = \beta_{ij}\mathbf{x}_j + \beta_{i\xi}\xi_j + c, \tag{2}$$

where we use the log of the continuous characteristics: MPG, lifetime miles, engine size, and empty weight.

A final assumption is that the choice set is continuous. Then, given an interior solution, the first

[14]The conditions are that (a) u_i is continuously differentiable in c and strictly increasing in c with $\frac{du}{dc} > \varepsilon$ for some $\varepsilon > 0$ and all $c \in (0, y_i)$; (b) u_i is Lipschitz continuous in $(\mathbf{x}_j, \xi_j)$; (c) u_i is strictly increasing in ξ_j.

[15]The unobserved characteristic may be influenced by marketing campaigns, for example.

order conditions imply

$$\beta_{ik} = x_{jk}\frac{d\tilde{p}}{dx_{jk}}, \tag{3}$$

where k indexes the four continuous characteristics. As detailed in section 4, the empirical challenge is to obtain estimates of β_{ik}, the random coefficients, by estimating the derivative of the price function. According to our model, β_{ik} represent truck owners' willingness-to-pay for a MPG, lifetime miles, engine size, and empty weight. The estimates of willingness-to-pay for MPG is a key component of our answer to whether truck owners undervalue the expected discounted gains from fuel efficiency.

3 Data

With the model and its assumptions in mind, in this section we describe the data. We first explain how we construct the data and then present summary statistics.

3.1 Origin of the data

The data we use is a compilation of two datasets, the Census Bureau's VIUS and the *Truck Blue Book*.[16] The VIUS was a survey conducted every five years in order to track the stock of trucks operating in the United States. (The survey was discontinued after 2002.) The Census Bureau surveyed the owners of a random sample of trucks registered or licensed in the United States as of July 1st of the survey year and recorded both physical and operational characteristics of the sampled truck. A few of the many characteristics in the VIUS are make and model-year of the truck, the vehicle identification number (VIN), fuel mileage, and lifetime miles. This survey, then, provides a detailed look at the stock of trucks every five years.

From the VIUS data on the stock of all trucks in 1992, we extracted a subset that fit our definition of long-haul trucks, or trucks designed for long-distance hauling. Based on conversations with industry analysts and a review of the trade press, we developed a list of criteria that trucks would

[16]Prior to 1997, the Census survey was known as the Truck Inventory and Use Survey. See `https://www.census.gov/svsd/www/vius/products.html` for more information. The *Truck Blue Book* is currently published by Primedia.

need to satisfy in order to be classified as a long-haul truck. First, we eliminated any truck that is not a truck-tractor. Truck-tractors are trucks designed to pull trailers, a necessary requirement for long-distance hauling. This restriction reduced the size of the sample from 123,641 observations to 42,108, as observations on pickup trucks and other medium trucks (e.g., straight trucks or 'box' trucks) were eliminated. This subset, however, still contained trucks that clearly were not used to haul goods over highways. For example, trucks that were extensively used off-road or had a body type incompatible with long-distance hauling (e.g., utility truck) remained in the sample. Consequently, we further refined the set of truck-tractors by only including those that satisfied the following criteria:

1. Have three axles
2. Have either a conventional or cabin-over-engine design
3. Have a diesel engine and air brakes
4. Do not spend most of their time off-road
5. Fit a list of body types[17]

The first restriction mainly eliminated truck-tractors with two axles; these trucks are limited by how much cargo they can pull, and they serve a niche market by hauling light loads. Similarly, this restriction ruled out trucks with four or more axles, which are a subset of trucks catering to a extremely small niche of firms typically engaged in 'severe service' activities. After conditioning on three axled truck-tractors, almost all trucks fulfilled criteria two and three. These constraints eliminated a few unusual trucks that are built to serve very particular demands. The fourth requirement was a check to make sure that the truck was operated in a manner consistent with long-haul trucks, whereas the last constraint ensured that the truck in question had a body type consistent with long-distance hauling. These restrictions slimmed down the dataset, decreasing the number of observations from 42,108 to 26,668.[18] We believe the resulting sample of observations is representative of the stock of class 8 long-haul trucks in the United States for the 1992 census year.

[17]The list of body types that were excluded from our refined sample were: pickup, panel or van, multistop or step van, garbage hauler, concrete mixer, yard tractor, sport utility, station wagon, minivan, and beverage, public utility, winch or crane, wrecker, service, oilfield and dump trucks.

[18]A summary of how we filtered the data is provided in table A1 in the appendix.

Although the VIUS provides a detailed accounting of the stock of long-haul trucks in the U.S., it does not provide the prices of these trucks. For this missing information, we turned to the *Truck Blue Book*, a comprehensive listing of used truck prices based on their characteristics. From the publisher, we obtained the October issues of the *Truck Blue Book* for 1992. We then assigned prices to trucks in the VIUS based on their recorded characteristics.

A difficulty in merging these two datasets is that the *Truck Blue Book* uses a different set of characteristics to describe a truck compared to the VIUS. Beyond more general characteristics such as make and model-year, it becomes difficult to distinguish trucks with different trim lines (e.g., differing engines or gross vehicle weight ratings). A solution to this problem is to use sections of the truck's vehicle identification number (VIN) as a link between the two data sources. The *Truck Blue Book* provides a portion of each truck's VIN. The VIUS collects, but does not publish the VIN. However, we were able to obtain a portion of the VINs for 1992 from the Census Bureau that revealed no confidential information.

The 1992 price data covers used trucks up to 8 years of age, which left us without prices for the oldest long-haul trucks. Using the sample weights, these older trucks accounted for 29 percent of the stock of long-haul trucks. Due to the lack of price data, we dropped these older trucks from our analysis. To merge the newer trucks with prices, we first aggregated the VIUS data to the make, model year, cabin type, and 2-digit VIN level. We then merged these data with the price data at this level of detail, resulting in a 75 percent match rate. For the trucks for which we could not find a match, we further aggregated up these observations to the make, model year, cabin type, and 1-digit VIN level, and matched them against our price data. This resulted in an additional 16 percent trucks being matched. The remainder of trucks were aggregated up to the make, model year, and cabin type level and matched to prices. In the end, the final dataset used for analysis has 629 observations.

Although we focus on 1992, we also merged the 1997 VIUS with the *Truck Blue Book* for October 1997 and used these data to check the robustness of our results. We merged these datasets using the method described above, in particular relying upon a 2-digit VIN acquired as a special tabulation from the Census Bureau.[19] The percent of observations matched using 2, 1 or no VIN digits for the 1997 data closely mimics what we observed for the 1992 data. The final 1997 dataset used for our robustness check has 510 observations.

[19] Details of the how the various filters reduced the size of the 1997 VIUS data are provided in table A1 in the Appendix.

3.2 Data description

This work focuses on seven truck characteristics: make, model-year, cabin type (i.e., conventional or cab-over-engine), engine size, weight when empty (a.k.a. empty weight), lifetime miles, and MPG. Using the VIUS, we identified the major brands that produce trucks-tractors in the United States, and we used industry information to consolidate brands that belonged to the same firm. In the end, we used six brands in our analysis: International, Kenworth, Mack, Peterbilt, Freightliner, and Ford. Freightliner is an agglomeration of brands including Freightliner, GMC/Chevy, White, and White GMC. Ford is also an agglomeration, including Ford, Autocar, Marmon, Scania, Volvo, and Western Star.

In the United States, conventional cabins are the dominant design, making up 74 percent of our sample. The cab-over-engine design was more popular in the past, and in our sample the fraction of cab-over-engine trucks increases with vintage; e.g., conventional trucks make up only 52 percent of vintage 8 trucks.[20]

Statistics describing the distribution of the 4 continuous characteristics are reported in table 1. Engine size is measured by the displacement of the engine in cubic inches, and survey respondents pick a rating bin which covers a range of 100 cubic inches. A rating of 18 corresponds to a displacement of 700 to 800 cubic inches, and the top range is 1001 cubic inches and above. (For comparison, the 2004 Honda Odyssey EX minivan had a displacement of just under 212 cubic inches.) The empty weight of a truck is measured in pounds, lifetime mileage is in miles, and MPG is the miles per gallon that the truck averaged in the survey year. As table 1 demonstrates, there is substantial variation across trucks in these four characteristics.

A central assumption of our model is that the product space is approximately continuous in the variables of interest, which are the four continuous characteristics. To provide a measure of how close trucks are in terms of their characteristics, we compute a nearest neighbor statistic. We then plot this statistic in figures 1 through 4 for each of the continuous characteristics. For MPG, the nearest neighbor is, on average, 0.0008 miles-per-gallon away. There are, of course, trucks on the edge of the product space that do not have close neighbors. For example, there were three trucks

[20]Cabin-over-engine designs were popular in the past because there existed regulations that restricted the length of the truck, where length was measured from the front bumper of the truck-tractor to back bumper of the trailer. By construction, the cabin-over-engine truck-tractors are shorter than the conventional cabins. In 1982, the federal government changed the regulation, and the length measurement focused on the trailer (and so excluded the truck-tractor).

Table 1: Long-haul truck characteristics

	Mean	SD	Percentiles 10th	25th	50th	75th	90th
MPG	5.69	4.46	5.00	5.30	5.63	5.97	6.32
Lifetime miles	393,332	174,846	120,000	223,931	366,873	507,375	609,519
Engine size	18.0	1.3	13.4	16.6	18.1	18.9	19
Empty weight	29,417	2,713	24,063	27,161	29,320	31,269	34,000

Note: MPG is miles per gallon and SD is standard deviation. The mean and standard deviation statistics were computed using the VIUS sample weights. Engine size is a displacement of the engine in cubic inches and is a categorical variable. A rating of 18 corresponds to a displacement of 700 to 800 cubic inches, and the top range is 1001 cubic inches and above. Empty weight is measured in pounds, and lifetime miles is in miles.

with less than 4 MPG, and a handful above 8 MPG. But these products are a tiny part of the sample and do not meaningfully affect our empirical results.

For lifetime miles, the median distance of the nearest-neighbor is 337 miles and for empty weight this statistic is 5 pounds. Both of these results suggest that the product space is approximately continuous in these characteristics. Of the four characteristics, engine size is potentially the most problematic because it is a categorical variable. The discrete nature of this variable is smoothed out, however, when aggregating trucks in the VIUS in order to be able to merge the price data (as described in section 3.1). Although there are bunching of trucks in the highest engine size, figure 3 displays what looks to be sufficient continuity along this characteristic. Indeed, the median nearest-neighbor statistic for engine size is 0.0008.

Turning to prices, we find substantial variation in prices across trucks. Indeed, there is a large decline in prices across vintages; vintage 0 long-haul trucks (i.e., trucks less than 1 year old) have a median price of $48,875 and vintage 8 trucks (i.e., trucks that are 8 years old) have a price of $8,412 (see figure 5). Further, within each vintage there is a wide range of prices, as illustrated by the large inter-quartile ranges displayed in figure 5. For example, the inter-quartile range is greater than $5,000 for vintage 0 trucks, which is greater than 10 percent of the median price.

Finally, to provide an overview of how these characteristics and price are related, we compute the correlations in the data among our continuous truck characteristics and price (see table 2). We find the correlations between price and each continuous variable are statistically significant. The lifetime miles and MPG are negatively correlated most likely because of technological progress; in our sample, average MPG increases as vintage decreases (i.e. as trucks get newer). The negative

Figure 1: Miles per gallon

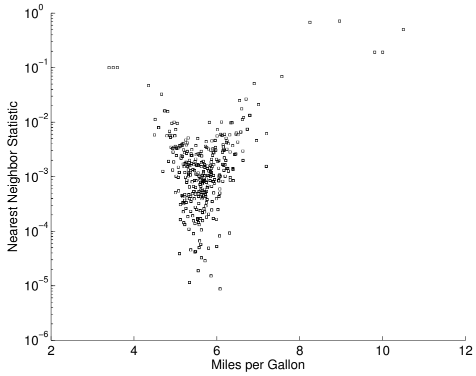


Figure 2: Lifetime miles

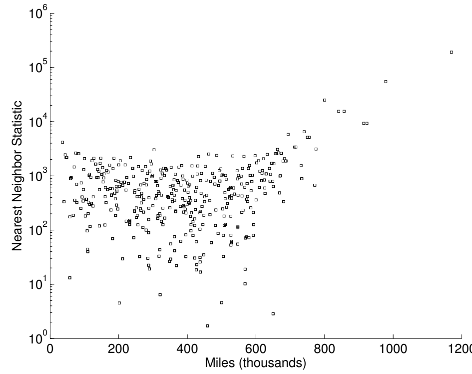


Figure 3: Engine size

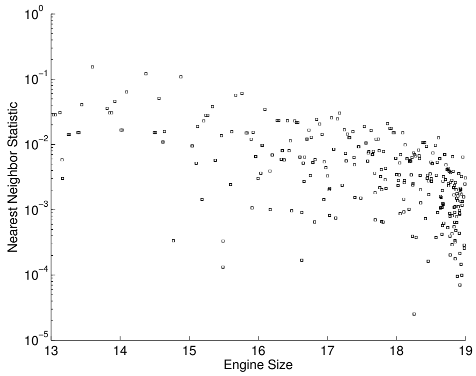
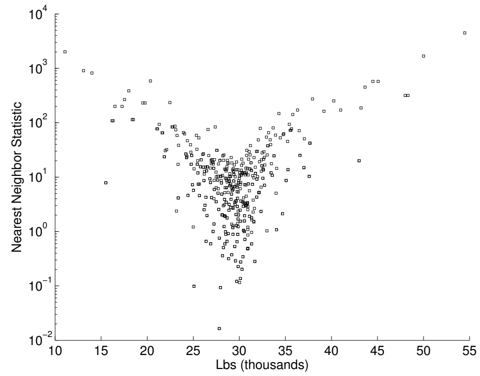


Figure 4: Empty weight



Note: For each truck, the distance to the nearest neighbor is computed for the four continuous characteristics. For each of these characteristics, this nearest-neighbor statistic is then plotted for each truck on a log-scale.

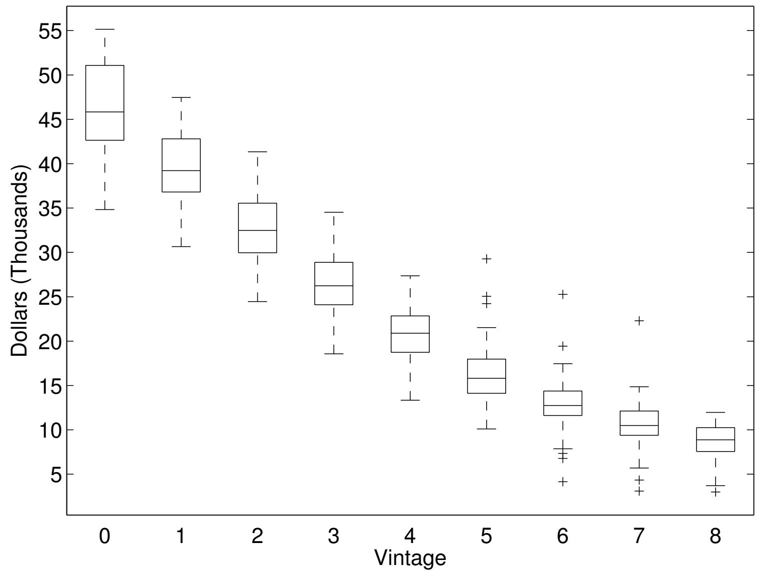
correlation between empty weight and MPG is likely driven by the fact making a truck lighter will increase its MPG, all else equal.

4 Empirics

4.1 Estimation

In order to recover truck owners' willingness to pay, we need to estimate the price hedonic and, more importantly, its derivatives (see equation 3). We accomplish this by using the data on truck prices and characteristics to estimate the conditional density of the price function, denoted $g(p|\mathbf{x}_j)$,

Figure 5: Price distribution by vintage



Note: A box-and-whiskers chart is displayed for each vintage. The upper, middle, and lower horizontal lines of the box portion correspond to the 75th, 50th, and 25th percentiles, respectively, of the price distribution of a given vintage. Let q_1 and q_3 denote the 25th and 75th quartile values, respectively. Then the upper most horizontal line, or whisker, is equal to $q_3 + 1.5(q_3 - q_1)$ and the lower most horizontal line is equal to $q_1 - 1.5(q_3 - q_1)$. Finally, the plus symbols denote outliers, or are all points which lie above the upper whisker or below the lower whisker.

Table 2: Correlations between variables

	Price	MPG	Lifetime miles	Engine Size	Empty weight
Price	1.00				
MPG	0.17^{***}	1.00			
Lifetime miles	-0.70^{***}	-0.16^{***}	1.00		
Engine size	-0.08^{*}	0.038	0.07	1.00	
Empty weight	0.11^{**}	-0.13^{***}	-0.005	0.003	1.00

Note: MPG is miles per gallon. Reported are Pearson correlation coefficients. The superscripts ***, **, * denote statistical significance at the 99, 95, and 90 percent confidence levels.

and its derivatives for every truck j in our sample. (Recall that p is price and $\mathbf{x}_j$ is a $K \times 1$ vector of characteristics.) To see the connection between the derivatives of the price hedonic and of the conditional density, note that

$$p(\mathbf{x}_j, \xi_j) = \int p g(p|\mathbf{x}_j) dp.$$

Then taking advantage of linearity, we have

$$\frac{\partial p_j(\mathbf{x_j}, \xi_j)}{\partial x_{j,k}} = \frac{\partial}{\partial x_{j,k}} \int p \cdot g(p|\mathbf{x}_j) dp, = \int p \cdot \frac{\partial g(p|\mathbf{x}_j)}{\partial x_{j,k}} dp, \tag{4}$$

where k denotes an element (a specific characteristic) in $\mathbf{x}_j$.

To impose as little parametric structure as possible, we estimate the conditional density of the price function and its derivatives using the local quadratic methods detailed in Fan and Gijbels (1996) and Fan et al. (1996). We can estimate $g(p|\mathbf{x}_j)$ using a nonparameteric regression technique, because as $h \to 0$,

$$E\left[K_h(p_j - p)|\mathbf{x}_j\right] \approx g(p|\mathbf{x}_j), \tag{5}$$

where K is a kernel density function, h is the bandwidth, and we define the scaled kernel $K_h(x) \equiv \frac{1}{h}K(x/h)$. Using Taylor's expansion, we know that for an x_0 in the neighborhood of x_j

$$E\left[K_h(p_j - p)|\mathbf{x}_j\right] \approx g(p|\mathbf{x}_j) + \frac{dg(p|\mathbf{x}_j)}{d\mathbf{x}_j}^T (\mathbf{x}_0 - \mathbf{x}_j) + \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_j)^T \mathbf{H}(\mathbf{x}_0 - \mathbf{x}_j), \tag{6}$$

where H is the Hessian matrix of $g(p|\mathbf{x}_j)$ with respect to $\mathbf{x}_j$, and the superscript T denotes the transpose. Fan et al. (1996) redefine the right hand side as

$$\alpha_j + \lambda_j^T(\mathbf{x}_0 - \mathbf{x}_j) + \gamma_j^T \text{vech}\left((\mathbf{x}_0 - \mathbf{x}_j)(\mathbf{x}_0 - \mathbf{x}_j)^T\right), \tag{7}$$

where vech(X) is the vectorization of the lower triangular portion of X and $\alpha_j \in \mathbb{R}, \lambda_j \in \mathbb{R}^K, \gamma_j \in \mathbb{R}^{K(K+1)/2}$. They then show that $(\alpha_j, \lambda_j, \gamma_j)$ can be estimated for every truck j by solving the weighted least squares problem

$$\min_{\alpha_j, \lambda_j, \gamma_j} \sum_{n=1}^{J} \left\{K_h(p_n - p) - \alpha_j - \lambda_j^T(\mathbf{x}_n - \mathbf{x}_j) - \gamma_j^T \text{vech}\left((\mathbf{x}_0 - \mathbf{x}_j)(\mathbf{x}_0 - \mathbf{x}_j)^T\right)\right\}^2 K_B(\mathbf{x}_n - \mathbf{x}_j), \tag{8}$$

where K_B is a multivariate kernel weighting function with bandwidth matrix B such that $K_B(\mathbf{u}) = (1/|B|)K(||B^{-1}\mathbf{u}||)$.[21] Our focus is on the price derivatives and so on the estimate of λ_j. Nevertheless, we include the higher-order terms in the minimization problem to reduce the bias of the estimator.[22] Further, if we wanted to recover estimates of the unobservable characteristics ξ_j for each truck, we could do so by computing the sample analog to equation (1), or

$$\hat{\xi}_j = \int_{p<p_j} g(p|\mathbf{x}_j)dp = \int_{p<p_j} \alpha_j(p)dp. \tag{9}$$

In our application, we use the univariate Gaussian kernel $K(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ to construct K_h and K_B, and take the bandwidth matrix to be of the form $B = h\mathbf{I}$. Since it is a scalar multiple of the identity matrix, this means that the smoothing is done along the coordinate axes, and that our estimator smooths by the same amount in every dimension. To ensure the variables are all on the same scale, we normalize each of them by its standard deviation for the purposes of computing the kernel weights.

The selection of the bandwidth parameter is a critical input of our chosen approach. We use least-squares cross-validation, a data-driven method where we choose a bandwidth that minimizes the out of sample prediction error over the sample space by estimating the model over sub-samples that leave out a single observation. Under this approach, the optimal bandwidth parameter is 2.48. More details on least-squares cross-validation can be found in appendix E.

Because the asymptotic properties of this price estimator do not depend on observing the individual firms in anything other than the cross section, we are able to recover estimates of the random coefficients from a single cross-sectional data set. Moreover, this estimator allows us to flexibly estimate the price hedonic and its derivatives without specifying a particular functional form. Given our focus on estimating the derivatives of the price hedonic at the observed prices, our method of estimating the conditional densities and integrating to recover the price hedonic is computationally equivalent to estimating the price hedonic directly. In appendix D, we provide more details on this point.

In this application, we use a log transformation of the price in order to estimate the price hedonic, and report willingness-to-pay as elasticities.

[21] Note that if $\mathbf{u} \in \mathbb{R}^1$, then this reduces to the scaled kernel K_h since B has only one entry.

[22] See Racine (2008) for a discussion of bias and nonparametric estimators. See Fan and Gijbels (1996) for a detailed discussion of bias and local polynominal estimators.

4.2 Results

We are able to recover the estimates of the willingness-to-pay for each truck owner in our sample for each of the continuous characteristics: MPG, lifetime miles, engine size, and empty weight. To generate an estimate of the distribution of willingness-to-pay in the population, we aggregate across truck owners using the sample weights in the VIUS. The kernel-smoothed distribution of willingness-to-pay for fuel efficiency, our main object of interest, is plotted in figure 6.[23] Our willingness-to-pay estimates are in terms of elasticities, so a coefficient of 0.6 for MPG means that a truck owner is willing to pay 0.6 percent of his truck's price for a 1 percent increase in MPG. Our estimates have the correct sign in that all, or almost all, truck owners value increases in MPG, engine size, and weight. Increasing MPG, or fuel efficiency, helps decrease the cost of operating a truck, and larger engine sizes make it easier to pull heavy loads. All else equal, heavier trucks are preferred because they are more comfortable to operate as, for example they typically vibrate and shake less. Further, almost all owners dislike lifetime miles, which provides us with the expected result that older trucks fetch lower prices.[24]

The distribution of tastes for fuel efficiency is quite wide, with the 10th and 90th percentiles of the distribution equal to 0.51 and 1.33 respectively (see table 3 and note that standard errors are bootstrapped, see appendix E for details). There is also substantial variety in tastes for lifetime mileage, as the 10th and 90th percentiles are -1.17 and -0.37, respectively.

Turning to engine size, we find that owners are quite sensitive to this measure, as they are willing to pay, on average, 1.39 percent of their truck's price for a 1 percent increase in engine size (i.e. cubic inch displacement). Further, there is wide dispersion in tastes among owners, with those in the left tail of the distribution hardly willing to pay for increased engine size (at the 10th percentile, our elasticity estimate is 0.14) whereas those in the right tail have elasticities greater than 2 (at the 90th percentile, our elasticity estimate is 2.72). In contrast, truck owners tastes for empty weight are more narrowly distributed, with an average elasticity of 0.45.

Returning to the fuel-efficiency estimates, we find that owners of newer vintage trucks have higher levels of willingness-to-pay for MPG. We illustrate this in figure 7, where we plot the cdf of

[23]For each graph involving a kernel-smoothed distribution, we computed the optimal bandwidth for the kernel-smoothing procedure using leave-one-out cross-validation to minimize the mean squared integrated error of the distribution (see appendix E for details).

[24]In appendix B are the figures illustrating the distribution of willingness-to-pay for lifetime miles, engine size, and empty weight.

Figure 6: Distribution of willingness-to-pay for miles per gallon

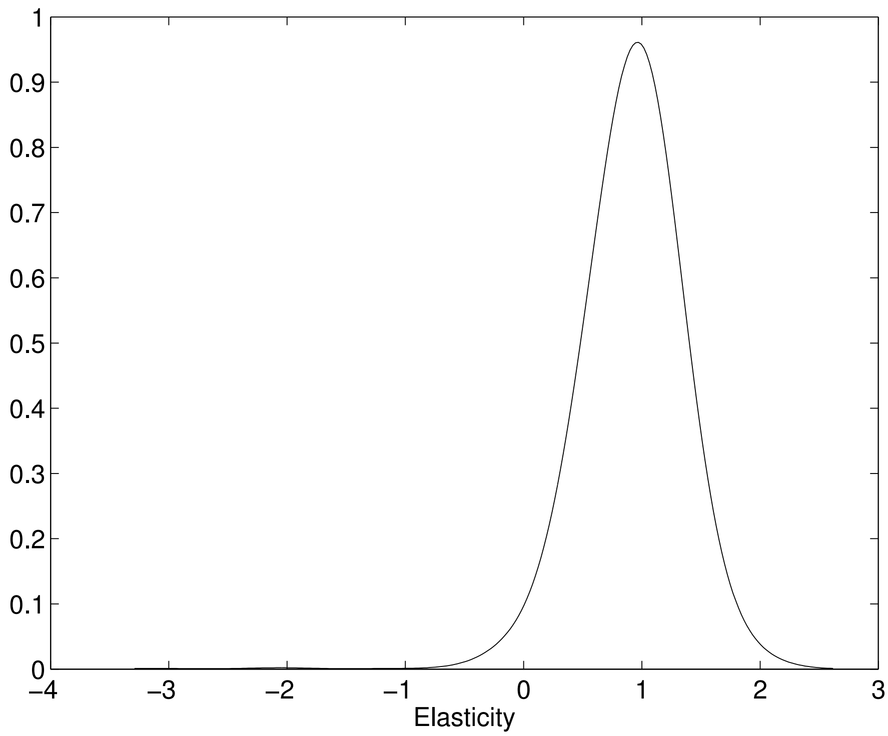
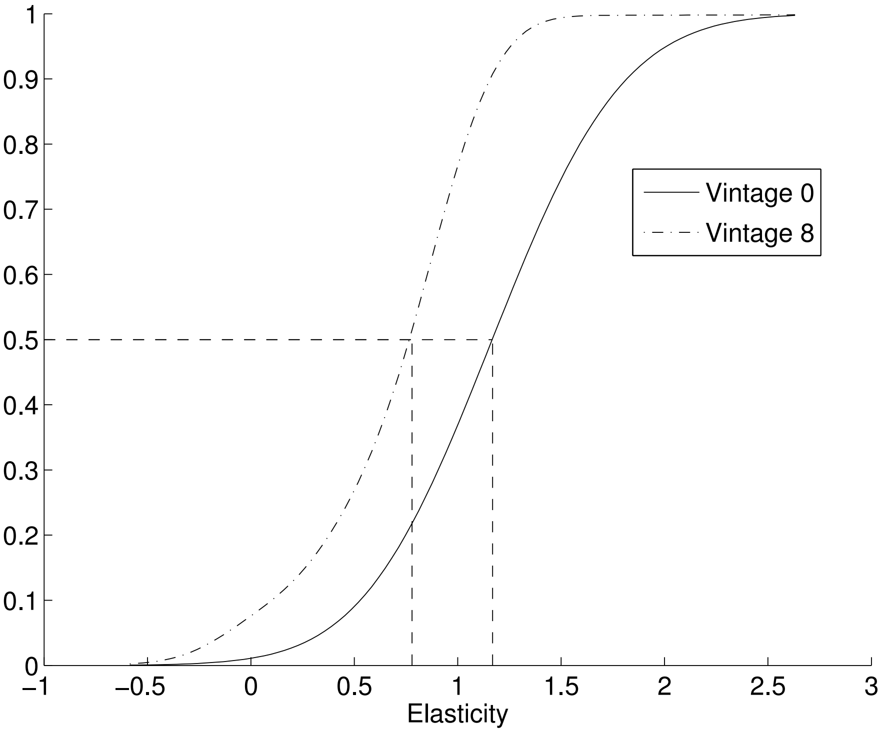


Table 3: Willingness-to-pay for truck characteristics (elasticities)

truck characteristic	mean	standard deviation	Percentiles 10th	25th	50th	75th	90th
Miles per Gallon	0.91	0.39	0.51	0.71	0.95	1.12	1.33
	(0.22)	(0.09)	(0.30)	(0.26)	(0.24)	(0.22)	(0.24)
Lifetime Miles	-0.85	0.30	-1.17	-1.06	-0.93	-0.68	-0.37
	(0.06)	(0.06)	(0.15)	(0.10)	(0.06)	(0.06)	(0.04)
Engine Size	1.39	1.05	0.14	0.82	1.45	2.04	2.72
	(0.46)	(0.26)	(0.27)	(0.37)	(0.54)	(0.66)	(0.73)
Empty Weight	0.45	0.25	0.24	0.34	0.49	0.60	0.69
	(0.14)	(0.06)	(0.18)	(0.16)	(0.15)	(0.15)	(0.17)

Note: Standard errors are in parenthesis and are computed by bootstrapping.

Figure 7: Distribution of willingness-to-pay for miles per gallon conditional on vintage purchased



willingness to pay for MPG conditional on owners having purchased a vintage 0 truck and a vintage 8 truck. From vintage 0 to vintage 8, there is a leftward shift in truckers' willingness to pay for MPG, evidenced by the median shifting from being greater than 1 for vintage 0 truck owners to less than 1 for vintage 8 truck owners.

5 Analysis

In this section, we compare our estimates of willingness to pay for MPG to the expected lifetime savings of an increase in fuel efficiency. For each truck j in our sample, we compute the expected discounted fuel savings associated with a 1 percent increase in MPG. These savings vary across trucks, because they are function of the truck's characteristics. Formally, we first define expected future fuel costs for a truck with mpg_j as

$$\text{FC}(\text{mpg}_j) = \sum_{v=\nu_j}^{L} \delta^{(v-\nu_j)} h(\nu_j, v) \left(\frac{m(v, x_j)}{\text{mpg}_j} \right) d(v - \nu_j)$$

where L is the maximum possible lifetime of a truck and δ is the discount rate. The function $h(\nu_j, v)$ is the probability that a truck of vintage ν_j survives to age v, $m(v, x_j)$ is the expected annual mileage of a truck of vintage v with characteristics x_j, and $d(x)$ is the expected price of diesel x years in the future from 1992. Expected fuel savings from a 1 percent increase in MPG is then just

$$\text{FS}(\text{mpg}_j) = \text{FC}(1.01 \cdot \text{mpg}_j) - \text{FC}(\text{mpg}_j). \tag{10}$$

To compute $\text{FS}(\text{mpg}_j)$ for every truck j, we need to determine (i) trucks' survival rates, (ii) expected annual mileage, (iii) the discount rate, and (iv) expected diesel prices.

5.1 Survival rates

We use the *Transportation Energy Data Book* (Davis et al., 2014) for the survival rates of heavy trucks and, based on this schedule, set the maximum age of a truck, L, to 30 years.[25] These survival rates are estimated using registration data on heavy trucks following the method outlined in Greenspan and Cohen (1996).

5.2 Expected annual mileage

To forecast trucks' expected mileage, we use the detailed VIUS data on class 8 long-haul trucks, which includes a variable recording annual mileage. This subset of the VIUS contains 26,668 observations and allows for a detailed analysis of how annual mileage varies with a truck's characteristics. In these data, truck age is recorded up to vintage 9; all trucks of vintage 10 and older are recorded with the same catch-all vintage 10. Further, we find that the annual mileage of the newest trucks, vintage 0, is unreliable. Annual mileage decreases with vintage except when comparing vintage 0

[25]The survival rates we use are reported in the Appendix, in table C2.

to vintage 1 trucks.[26] This unusual feature probably reflects the fact that vintage 0 truck owners are more likely to own (and so drive) their trucks for less than a year, which makes them more likely to understate their annual mileage. As such, in our analysis we drop vintage 0 and vintage 10 observations.

We use a regression to estimate how annual mileage varies with truck characteristics. In addition to fitting the data well, a goal of this approach is to forecast annual mileage out of sample, because our expected fuel savings measure takes into account the potential for a truck to be operated for 30 years. We found that a log-log specification performed best. In contrast, the linear and quadratic specifications we tried, while providing better in-sample fits, consistently predicted negative annual mileage out-of-sample. Formally, we estimate the following regression using weighted least-squares:

$$\begin{aligned}\log(\text{miles}_j) &= \theta_0 + \theta_1 \nu_j + \theta_2 \log(\text{mpg}_j) \\ &\quad + \theta_3 \log(\text{engsize}_j) + \theta_4 \log(\text{empwt}_j) + \theta_5 \text{cabtype}_j + \sum_{k=1}^{5} \kappa_k 1_{\text{make}_j = k} + \varepsilon_j, \quad (11)\end{aligned}$$

where miles_j is annual miles, engsize_j is engine size, empwt_j is empty weight, cabtype_j is a dummy variable equal to 1 for trucks with a conventional cabin type, make_j is the brand of truck j, and ε_j is an error term. The variable $1_{x=y}$ is an indicator function equal to 1 when $x = y$, and so controls for brand differences in the above regression. The weights used when estimating the regression are the VIUS sample weights.

We find the expected result that vintage has a negative impact on annual mileage (see table 4). Further, there are large differences across brands in terms of mileage. More surprisingly, we find that the effect of MPG on miles driven, although estimated to be a small positive number, is statistically insignificant.

The log-log specification provides a reasonable forecast of annual mileage out-of-sample, for vintages 10 through 30. As a rough check on these predictions, we compare them against the annual mileage reported by the old trucks in the our data (those trucks for which we only know that the vintage is greater than 9). Specifically, for these trucks we compute the 10th, 25th, 50th, 75th, and 90th percentiles of the distribution of annual miles. We then predict annual miles out-

[26] Vintage 0 are 1993 and 1992 model year trucks. Vintage 1 are 1991 model year trucks. This same odd increase in annual mileage from vintage 0 to vintage 1 is observed in the 1997 VIUS.

Table 4: Predicting annual mileage, estimated coefficients

Independent Variable	Coefficient	
	Estimate	SE
Intercept	10.56	0.17
vintage	-0.10	0.00
log(mpg)	0.04	0.03
log(engsize)	0.23	0.05
log(empwt)	0.07	0.02
cabtype	-0.11	0.01
International brand	0.23	0.02
Kenworth brand	0.37	0.02
Mack brand	0.02	0.03
Peterbuilt brand	0.39	0.03
Freightliner brand	0.26	0.02

Note: Reported are the estimated coefficients of a regression where the dependent variable is the log of annual mileage. SE is standard error and Ford is the reference brand. There are 16,864 observations and the R-squared is 0.13.

of-sample for all (the young) trucks in the data used to estimate the regression, supposing these trucks' vintage is equal to 10, 11, ... , and 30. We then compute the percentiles of the resulting distribution of predicted annual miles.[27] We emphasize this is a rough check of our out-of-sample predictions, because truck characteristics are likely to be different across the sets of young and old trucks. Reassuringly, the distribution of predicted miles is somewhat close to the distribution of actual miles, at least for the 25th, 50th and 75th percentiles (see table 5). We have confidence then, that our prediction of annual mileage out-of-sample are reasonable.

5.3 Discount rate

We assume that truck owners' discount rate is 6 percent. We arrive at this rate by first taking Moody's BAA rate as indicative of the rates that owner's receive when seeking to finance the purchase a truck.[28] In 1992, the average nominal rate for a BAA corporate bond was 8.98 percent. This

[27]To account for the fact that trucks might be scrapped before reaching vintage 30 we use the survival rates described earlier in this section as weights.

[28]BAA indicates there is moderate credit risk.

Table 5: Out-of-sample check on predicted annual mileage (vintages 10+)

	Percentiles				
	10th	25th	50th	75th	90th
Data	3,000	8,000	27,417	52,503	80,000
Predicted	9,755	15,150	24,506	34,832	43,375

Note: Data are trucks with a vintage greater than 9 in the detailed VIUS data set of long-haul trucks. The percentiles for these trucks' annual mileage are computed using VIUS sample weights. Predicted are an out-of-sample forecast of annual miles, given vintages equal to 10, 11, ... , 30. The percentiles of the predicted annual miles are computed using weights equal to the product of the VIUS sample weights and survival rates.

rate seems reasonable given the average rate in 1992 on 48 month new car loans to households was 9.3 percent.[29] Rounding to 9 percent and accounting for the 3 percent rate of inflation in 1992, we end up with discount rate of 6 percent.[30]

5.4 Expected diesel prices

Finally, we assume that truck owners view diesel prices as a random walk, and therefore expect future prices to be equal to the current price.[31] Based on a combination of World Bank and Energy Information Administration (EIA) data on prices, we compute that the average price of diesel per gallon in the U.S. in 1992 is $1.084.[32] In the robustness section, we re-do our analysis assuming that truck owners have perfect foresight over diesel prices and find that our main results do not change.

5.5 Expected fuel savings

From these data, we compute the expected fuel savings corresponding to a 1 percent increase in fuel efficiency. In our sample, the expected fuel savings vary considerably, from a minimum of $225 to

[29] This rate is published by the Board of Governors in their G.19 Consumer Credit report.

[30] We use the Bureau of Labor Statistics consumer price index for all urban wage earners to calculate the rate of inflation in 1992.

[31] Busse et al. (2013) model consumers' expectations of gasoline prices as following a random walk, based on evidence presented in Anderson et al. (2011).

[32] The World Bank publishes U.S.diesel prices per liter, which we converted to gallons. The (converted) prices are $1.060 and $1.022 for 1992 and 1998 respectively. The EIA's published price in 1998 is $1.044. Applying the percent change in World Bank prices to the EIA price, we arrive at a diesel price of $1.084 in 1992.

Table 6: Expected future fuel savings by vintage and MPG (dollars)

Vintage	5 MPG	6 MPG	7 MPG	8 MPG
0	1,341	1,125	970	853
1	1,193	1,001	863	759
2	1,058	888	766	673
3	934	784	676	594
4	820	688	593	522
5	728	611	526	463
6	646	542	467	411
7	574	481	415	365
8	510	428	369	324

Note: MPG is miles per gallon. This table reports the expected future fuel savings of a 1 percent increase in MPG for a given pair of vintage and MPG, holding all other truck characteristics fixed at their mean values.

a maximum of $1,522. However, 80 percent of truck owners fall between $489 and $1,121, the 10th and 90th percentiles respectively. To provide a sense of how these savings depend on the MPG and age of each truck, in table 6 we report our calculations of the expected lifetime savings for specific vintage and MPG pairs, holding fixed all other truck characteristics at their mean values.

Owners of lower MPG trucks gain the most from an increase in fuel efficiency, because their total fuel costs are appreciably higher. Similarly, owners of younger trucks can expect higher fuel savings because they can expect more years of usage over which to accumulate the savings.[33]

We now combine our calculations of net present value of future fuel savings with our estimates of willingness to pay. We do this by constructing a ratio were the numerator is a truck owner *j*'s WTP for a 1 percent increase in MPG and the denominator is the associated net present value of future fuel savings for a 1 percent increase in MPG. If this ratio is equal to 1, then a truck owner's WTP for future fuel savings is equal to the expected benefits, whereas if this ratio is less than 1, the truck owner undervalues future fuel savings. We convert this measure into a percent, aggregate across truck owners using the VIUS sample weights, and graph its (kernel-smoothed) distribution in figure 8.

[33]In our analysis, we abstract away from tax considerations. This is because trucking corporations can treat both fuel costs and fuel-efficiency investments as expenses when calculating their net income for tax purposes. Similarly owner-operators can deduct both fuel costs and fuel-efficiency investments from their taxes.

Figure 8: Distribution of truck owners' valuation of lifetime fuel savings

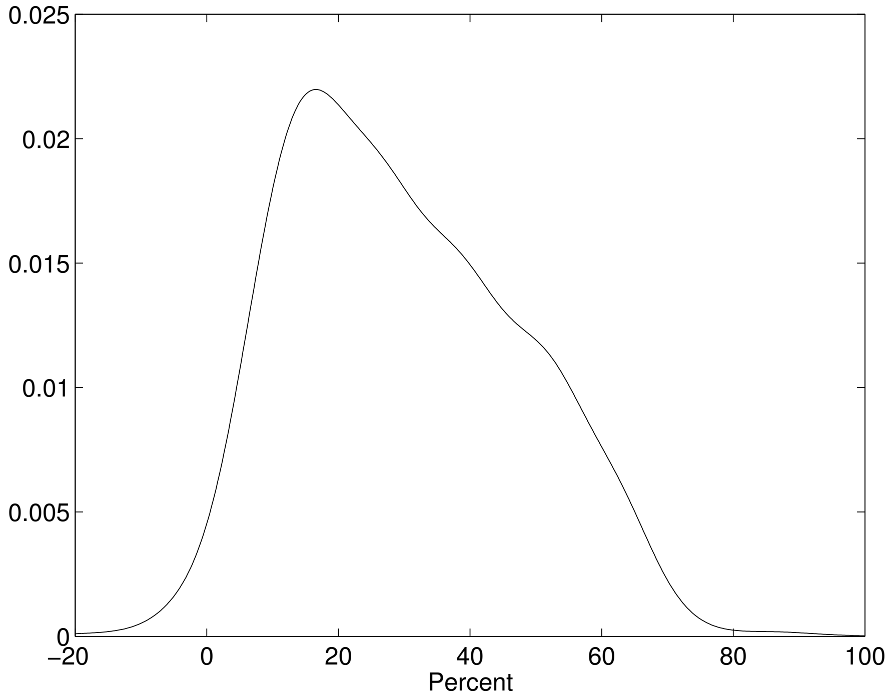


Table 7: Summary statistics on truck owners' valuation of lifetime fuel savings (percent)

mean	standard deviation	Percentiles 10th	25th	50th	75th	90th
29.5	19.3	8.8	15.8	27.1	41.9	54.5
(0.1)	(0.1)	(0.2)	(0.2)	(0.1)	(0.1)	(0.2)

Note: Standard errors are in parenthesis and are computed using bootstrapping.

As illustrated in figure 8, truck owners considerably undervalue lifetime fuel efficiency savings. On average, owners are willing to pay for only 29.5 percent of the lifetime fuel cost savings that accrue from a 1 percent increase in MPG. At the 90th percentile, owners are willing to pay for 54.5 percent of the lifetime fuel savings and those at the 10th percentile are willing to pay for only 8.8 percent the savings (see table 7).

Another way to view these results is to compute the discount rate at which a truck owner's willingness to pay for a 1 percent increase in MPG is equal to the associated lifetime fuel efficiency savings. We calculate these discount rates across all long-haul truck owners and find that it ranges from 0.31 to 0.82 (the 10th and 90th percentile, respectively) with a median value of 0.64, which is markedly higher than the 0.06 that we assumed in section 5.3.

The heterogeneity among truck owners with regard to their willingness to pay for fuel efficiency is related to their truck's vintage. In particular, owners of newer trucks have a higher valuation of future fuel savings than owners of older trucks. We illustrate this point by plotting the cdf of the distribution of owner's valuation of future fuel savings conditional on the owners truck being of vintage 0 and vintage 8 (figure 9).

In comparing these two cdfs, we see that the median vintage 0 truck owner is willing to pay for about 53 percent of future fuel savings, whereas the median vintage 8 owners is willing to pay about 9.7 percent. This pattern is reported in Allcott and Wozny (2014) for light vehicles, and likely reflects usage, in that the average annual mileage of long-haul trucks steadily decreases by vintage. Whereas vintage 0 trucks are driven more than 100,000 miles in a year, vintage 8 trucks are driven about 63,000 miles. Because of the difference in willingness-to-pay, the model predicts that technologies that improve fuel efficiency will be disproportionately incorporated on the newest trucks.

Another difference across truck owners is that those which own more fuel-efficient trucks have a

Figure 9: Cdf of vintage 0 and vintage 8 truck owners' valuation of future fuel savings

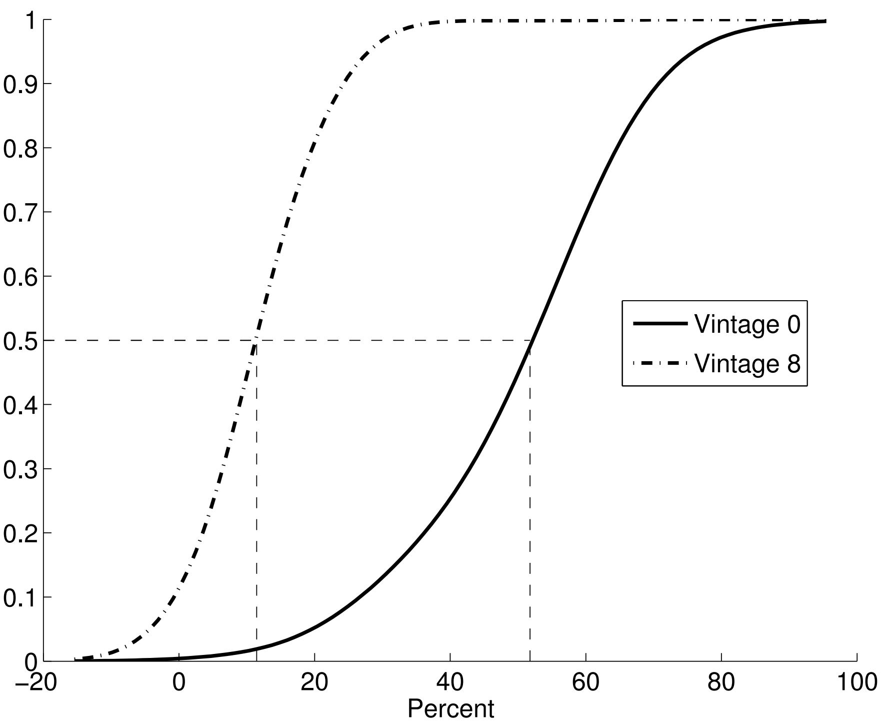
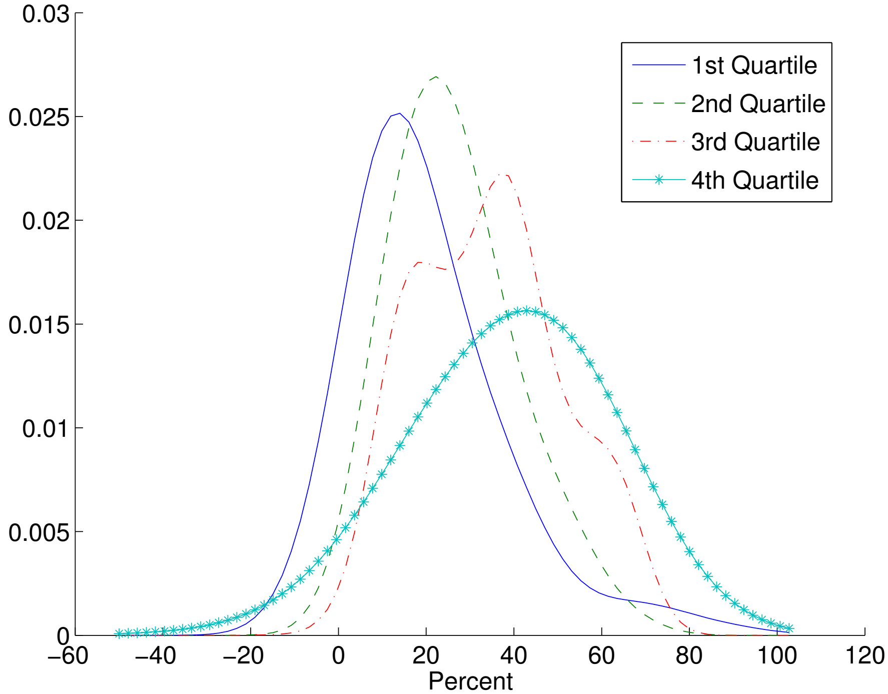


Figure 10: Truck owners' valuation of future fuel savings by MPG quartile



higher valuation of future fuel savings. To demonstrate this feature of our results, we divide the set of truck owners into quartiles, based on the MPG of their truck-tractor. We then plot the distribution of truckers' fuel savings valuation for each quartile in figure 10.[34] As illustrated in the figure, long-haul truckers who purchased high MPG trucks have fuel savings valuation closer to 100 percent relative to those which purchased low MPG trucks.

5.6 Robustness

In this section, we consider the robustness of our main result along two dimensions. We first repeat our analysis using the 1997 VIUS and October 1997 *Truck Blue Book* price data. As described at the end of section 3, we are able to merge these two data sets using the same approach taken for the

[34] Rather than showing a strict decomposition of the probability distribution of the valuation ratio, we explicitly re-estimate the distribution of the valuation ratio for each quartile. As a result, the optimal bandwidths may differ from the bandwidth chosen for the whole population. This explains why we observe nonzero mass above 80% in the distribution of some of the quartiles, but not in the distribution for the whole population.

Table 8: Distribution of undervaluation using different assumptions

	Percentiles					Mean	SD
	10th	25th	50th	75th	90th		
1992 VIUS	8.8	15.8	27.1	41.9	54.5	29.5	19.3
Benchmark	(0.2)	(0.2)	(0.1)	(0.1)	(0.2)	(0.1)	(0.1)
1997 VIUS	2.4	7.5	10.8	15.8	22.9	11.8	8.5
	(0.2)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Perfect foresight	8.4	15.0	25.6	39.2	51.1	27.7	18.0
	(0.2)	(0.2)	(0.1)	(0.1)	(0.2)	(0.1)	(0.1)

Note: Benchmark results are the same as those reported in table 7. 1997 VIUS are the results when using the 1997 VIUS and October 1997 *Truck Blue Book* prices. Perfect foresight are the results when using realized real diesel prices from 1992 onward. SD is standard deviation. Standard errors are in parenthesis and are computed using bootstrapping.

1992 data. In repeating our work on the 1997 data sets, we find that long-haul truckers substantially undervalue future fuel savings, confirming the benchmark 1992 results. In fact, the undervaluation is larger in 1997, with the average long-haul trucker willing to pay for only 11.8 percent of the expected future fuel savings associated with a 1 percent increase in MPG (see the second row of table 8).

We then checked the robustness of the assumption that truckers' view diesel prices as a random walk by recomputing the expected future fuel savings using actual diesel prices. Annual on-highway diesel prices are published by the Energy Information Administration from 1995 onward (see table C3 in the appendix). We linearly interpolate prices for 1993 and 1994 and assume a random walk for diesel prices after 2014.[35] Under this assumption, our main undervalution result still holds, with the average long-haul trucker willing-to-pay for only 27.7 percent of the expected future fuel savings from a 1 percent increase in MPG.

5.7 SuperTruck analysis

In the 2011 final ruling announcing the fuel efficiency regulations, the NHTSA and the EPA emphasized that heavy truck manufacturers should be able to meet the new fuel efficiency requirements using already existing technologies. Innovations such as low-resistence tires, better aerodynam-

[35]Because of discounting, how diesel prices are extrapolated beyond 2014 has little effect on our results.

ics, and incremental improvements in heavy duty engines can all be used, these agencies argue, to dramatically improve fuel efficiency.[36]

The SuperTruck program aptly makes this argument. This research program was funded by the Argonne National Laboratory and the Department of Energy with the goal of demonstrating the feasibilty of a 50 percent improvement in fuel efficiency for class 8 long-haul trucks compared with current models.[37] As part of this research program, four teams headed by different truck manufacturers modified a long-haul truck with the goal of dramatically increasing its fuel efficiency. The result from this work are summarized with the presentation of two different technology platforms. The first increases the truck's MPG by 65.3 percent for a cost equal to 26.6 percent of the price of a new truck. The second modification increases MPG by 69.8 percent for a cost equal to 51 percent of a new truck. The report than makes predictions about truck owners' willingness to adopt these new technologies over time.

Given the fuel efficiency benefits and costs of the two SuperTruck technology platforms, our estimates of willingness-to-pay provide insight on whether truck owners would adopt either platform. In particular, using our estimated elasticities, we can compute the fraction of vintage 0 truck owners that would be willing to pay these costs in order to benefit from the higher fuel efficiency. For the first SuperTruck platform, a truck owner with an elasticity of 0.407=26.6/65.3 would be indifferent between adopting or not adopting the technological improvements. For the second modification, the indifferent truck owner has an elasticity of 0.731. Based on our estimated distribution of willingness to pay for fuel efficiency for vintage 0 truck owners (see figure 7), we find that 93.6 percent of new heavy truck owners would be willing to pay the costs of the first modification for the increase fuel efficiency. The second modification would be slightly less popular, with 80.9 percent of new heavy truck owners willing to pay for the costs associated with the increase in fuel efficiency. Further, using a compensating variation approach, we find that all truck owners strictly prefer the first modification over the second.

Our results are encouraging in that the model predicts that a strong majority of new truck owners are willing to bear the costs of adopting the fuel-saving innovations proposed by the SuperTruck program. A caveat with our analysis is our estimates are based on observations recorded in 1992, more than twenty years ago. Although we do not expect these deep parameters of the truck owner's

[36]For detailed arguments, see the Regulatory Impact Analysis report on these regulations jointly published by the EPA and the NHTSA at http://www.nhtsa.gov/staticfiles/rulemaking/pdf/cafe/Truck_CAFE-GHG_RIA.pdf.

[37]The final report of this program can be found at http://www.transportation.anl.gov/pdfs/TA/903.PDF.

problem to vary much with time, it is not unreasonable to worry that these elasticities may have changed over twenty years.

6 Conclusion

In this paper, we estimate that truck owners of long-haul trucks willingness-to-pay for MPG. We find there is a wide range in willingness-to-pay across truck owners, with the elasticity of fuel efficiency to price ranging from 0.51 (10th percentile) to 1.33 (90th percentile). On average, trucks owners are willing to pay 0.91 percent of their truck price for a one percent increase in MPG. We then compute the expected lifetime savings from a 1 percent increase in MPG and compare this measure against truck owners' willingness-to-pay. Overall, we find that truck owners undervalue future fuel savings from increased fuel efficiency; on average, owners are willing to pay for only 29.5 percent of the expected lifetime savings associated with a 1 percent increase in MPG. This low valuation of fuel efficiency suggests that the federal government's policy of setting fuel efficiency standards for medium and heavy trucks could be an effective policy tool to raise fuel economy.

Looking ahead, a substantial amount of research still needs to be done analyzing the effectiveness of fuel standards on medium and heavy trucks. We are interested to see whether subsequent studies on long-haul truck-tractors, which presumably will find and use different datasets and techniques, will support our main results. Furthermore, more work needs to be done determining why truck owners' undervalue future fuel savings (e.g., Vernon and Meier (2012) explore how principal-agent problems can lead to under-investment in fuel efficiency technologies). Finally, the overall effectiveness of fuel efficiency standards depends upon their effects on more than the subset of medium and heavy trucks that we considered in our analysis.

References

Ackerberg, Daniel, Lanier Benkard, Steven Berry, and Ariel Pakes (2007) "Econometric Tools for Analyzing Market Outcomes," in James Heckman and Edward Leamer eds. *Handbook of Econometrics*: North-Holland.

Allcott, Hunt and Nathan Wozny (2014) "Gasoline Prices, Fuel Economy, and the Energy Paradox," *Review of Economics and Statistics*, Vol. 96(5), pp. 779–796.

Anderson, Soren T., Ryan Kellogg, and James M. Sallee (2011) "What Do Consumers Believe about Future Gasoline Prices?", NBER 16974.

Bajari, Patrick and Lanier Benkard (2005) "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy*, Vol. 113(6), pp. 1239–1276.

Bajari, Patrick and Matthew Kahn (2005) "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities," *Journal of Business and Economic Statistics*, Vol. 23(1), pp. 20–33.

Baker, George and Thomas Hubbard (2004) "Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking," *Quarterly Journal of Economics*, Vol. 119(4), pp. 1443–1479.

Bento, Antonio, Shanjun Li, and Kevin Roth (2012) "Is There an Energy Paradox in Fuel Economy? A Note on the Role of Consumer Heterogeneity and Sorting Bias," *Economic Letters*, Vol. 155(1), pp. 44–48.

Berry, Steve, James Levinsohn, and Ariel Pakes (1995) "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63(4), pp. 841–890.

Busse, Meghan, Christopher Knittel, and Florian Zettelmeyer (2013) "Are Consumers Myopic? Evidence from New and Used Car Purchases," *American Economic Review*, Vol. 103(1), pp. 220–256.

Davis, Stacy, Susan Diegel, and Robert Boundy (2014) "Transportation Energy Data Book," Prepared for the U.S. Department of Energy.

Fan, J. and I. Gijbels (1996) *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability 66*.

Fan, Jianqing, Qiwei Yao, and Howell Tong (1996) "Estimation of Conditional Densities and Sensitivity Measures in Nonlinear Dynamical Systems," *Biometrika*, Vol. 83, pp. 189–206.

Goldberg, Pinelopi (1998) "The Effects on the Corporate Average Fuel Economy Standards in the US," *Journal of Industrial Economics*, Vol. 46, pp. 1–33.

Greenspan, A. and D. Cohen (1996) "Motor Vehicle Stocks, Scrappage, and Sales," Finance and economics discussion series, v96(40).

Grigolon, Laura, Mathias Reynaert, and Frank Verboven (2014) "Consumer Valuation of Fuel Costs and the Effectiveness of Tax Policy: Evidence from the European Car Market," University of Leuven working paper.

Helfand, Gloria and Ann Wolverton (2011) "Evaluating the Consumer Response to Fuel Economy : A Review of the Literature," *International Review of Environmental and Resource Economics*, Vol. 5, pp. 103–146.

Hubbard, Thomas (2003) "Information, Decisions, and Productivity: On Board Computers and Capacity Utilization in Trucking," *American Economic Review*, Vol. 93(4), pp. 1328–1353.

Koster, Hans, Jos van Ommeren, and Piet Rietveld (2014) "Agglomeration Economies and Productivity: A Structural Estimation Approach Using Commercial Rents," *Economica*, Vol. 81, pp. 63–85.

Leard, Benjamin, Joshua Linn, Virginia McConnell, and William Raich (2015) "Fuel Costs, Economic Activity, and the Rebound Effect for Heavy-Duty Trucks," Sept, RFF Discussion Paper 15-43.

Maclean Hunter Market Reports, Inc. (1992a) *1992 Truck Identification Book: Official Vehicle Identification*, Vol. 18, No. 1, Chicago, IL.

——— (1992b) *The Truck Blue Book*, Vol. 82, No. 4, Chicago, IL.

National Market Reports, Inc. (1997a) *1997 Truck Identification Book: Official Vehicle Identification*, Vol. 23, No. 1, Chicago, IL.

——— (1997b) *The Truck Blue Book*, Vol. 87, No. 4, Chicago, IL.

NHTSA-EPA (2011) "Greenhouse Gas Emissions Standards and Fuel Efficiency Standards for Medium- and Heavy-Duty Engines and Vehicles," Published in the Federal Register [vol 76(179)] on September 15, 2011.

——— (2015) "Proposed Rulemaking for Greenhouse Gas Emissions and Fuel Efficiency Standards for Medium- and Heavy-Duty Engines and Vehicles – Phase 2," Draft Regulatory Impact Analysis.

Parry, Ian, Margaret Walls, and Winston Harrington (2007) "Automobile Externalities and Policies," *Journal of Economic Literature*, Vol. 45, pp. 373–399.

Racine, Jeffrey (2008) "Nonparametric Econometrics: A Primer," *Foundations and Trends in Econometrics*, Vol. 3(1), pp. 1–88.

Thomassen, Oyvind (2014) "A Pure Characteristics Demand Model for Automobile Variants," *Seoul Journal of Economics*, Vol. 27, pp. 187–222.

Torrey IV, W. Ford and Dan Murray (2014) "An Analysis of the Operational Costs of Trucking: 2014 Update."

Vernon, David and Alan Meier (2012) "Identification and Quantification of Principal–Agent Problems Affecting Energy Efficiency Investments and Use Decisions in the Trucking Industry," *Energy Policy*, Vol. 49, pp. 266–273.

West, Sarah and Roberton Williams III (2005) "The Cost of Reducing Gasoline Consumption," *American Economic Review Papers and Proceedings*, Vol. 95(2), pp. 294–99.

Wollmann, Thomas (2014) "Trucks without Bailouts: Equilibrium Product Characteristics for commercial vehicles," Dec, Harvard University manuscript.

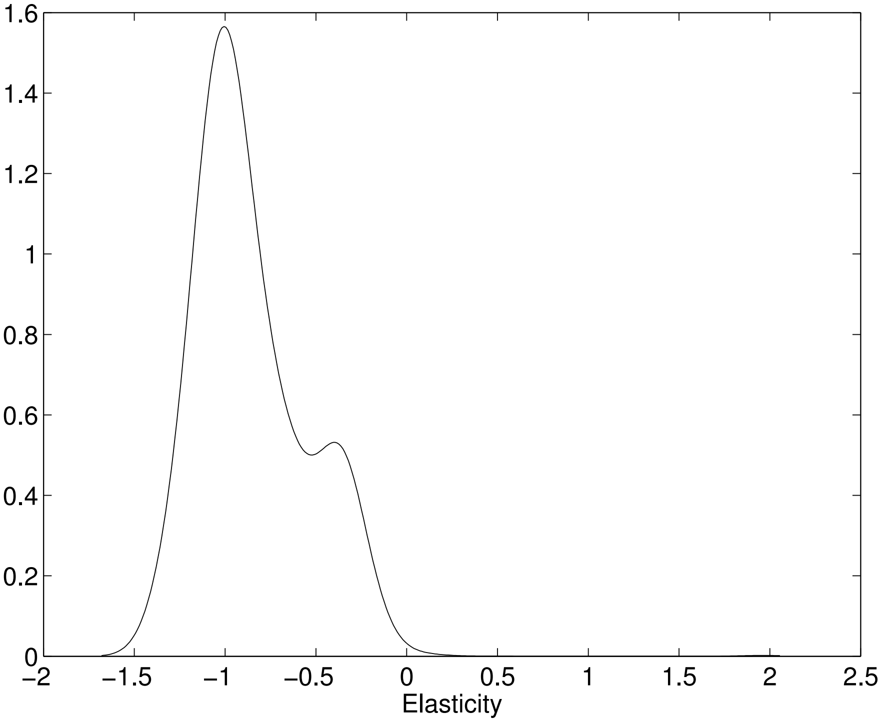
Table A1: Filtering of the 1992 and 1997 VIUS data

	Observations		Sample weights	
	number	percent	sum	percent
1992				
VIUS data set	123,641	100.0	5,920,075,519	100.00
Only truck-tractors	42,108	34.1	116,817,124	1.97
Have three axles	32,240	26.1	83,909,899	1.42
Have a conventional or cabin-over-engine design	31,335	25.3	81,696,386	1.38
Have diesel engine and air brakes	30,591	24.7	79,208,583	1.34
Do not spend most of their time off-road	29,588	23.9	77,244,933	1.30
Has the correct body type	26,668	21.6	70,086,661	1.18
1997				
VIUS data set	104,545	100.0	72,800,251,891	100.00
Only truck-tractors	27,956	26.7	1,543,752,184	2.12
Have three axles	21,749	20.8	1,140,559,205	1.57
Have a conventional or cabin-over-engine design	20,611	19.7	1,074,940,667	1.48
Have diesel engine and air brakes	19,867	19.0	1,030,056,985	1.41
Do not spend most of their time off-road	19,292	18.5	1,005,571,131	1.38
Has the correct body type	17,385	16.6	919,831,799	1.26

A Filtering of the VIUS data

In table A1 we report how the different filters we used to identify class 8 long-haul trucks in the VIUS reduced the size of the sample, both in terms of the number of actual observations and the implied aggregate number of trucks based upon the survey's sample weights.

Figure B1: Distribution of willingness-to-pay for lifetime miles



B Willingness-to-pay figures

Figures B1 to B3 illustrate the kernel-smoothed distribution of long-haul truckers willingness-to-pay for lifetime miles, engine size, and empty weight, respectively. As noted in the paper, for each graph, we computed the optimal bandwidth for the kernel-smoothing procedure using leave-one-out cross-validation to minimize the mean squared integrated error of the distribution (see appendix E for details).

Figure B2: Distribution of willingness-to-pay for engine size

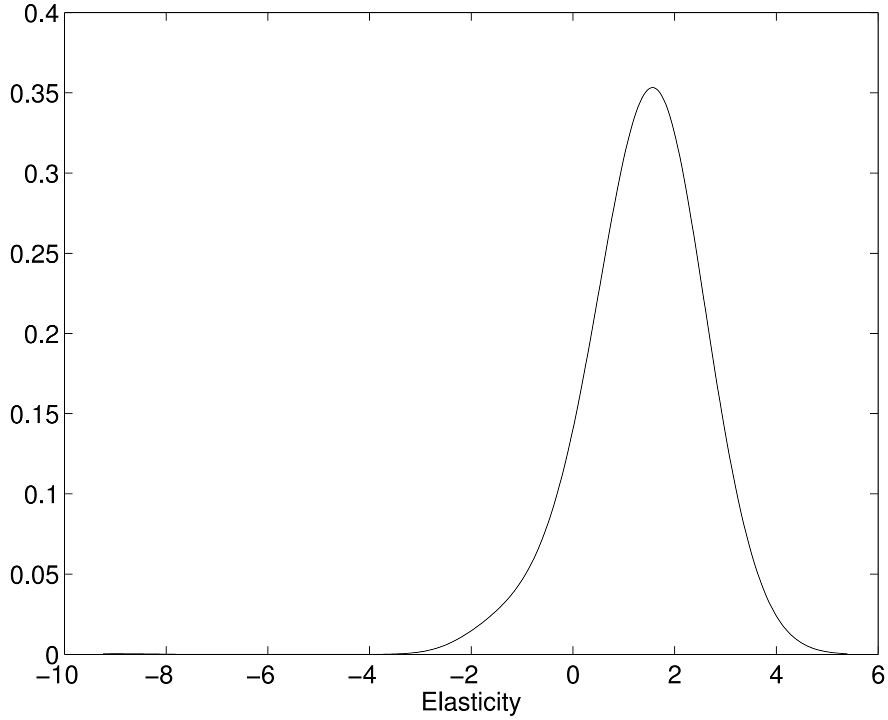


Figure B3: Distribution of willingness-to-pay for empty weight

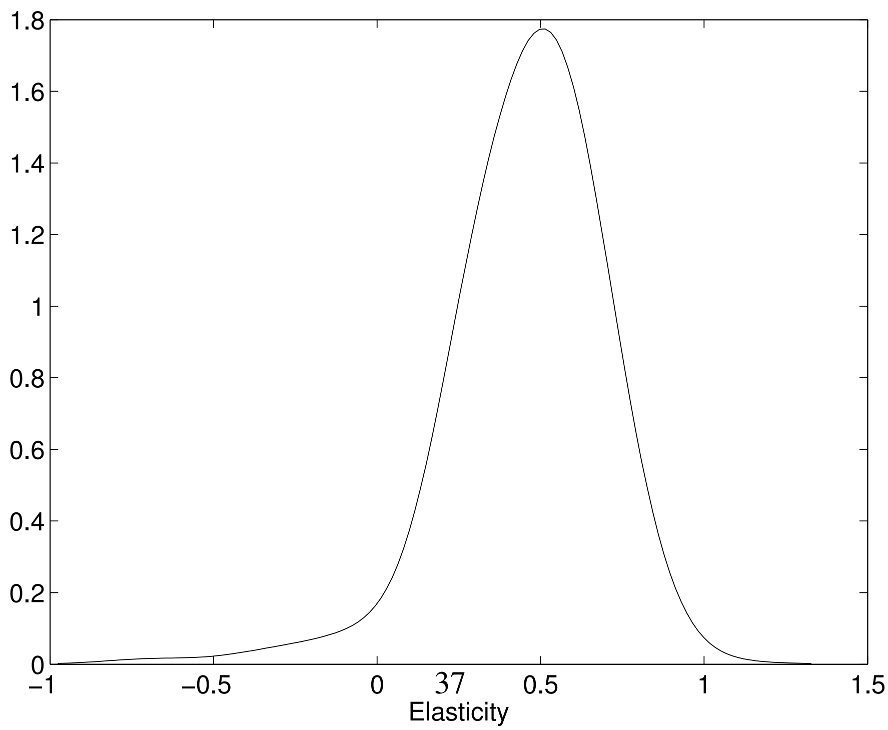


Table C2: Survival rates by vintage

Vintage	1	2	3	4	5	6	7	8	9	10
Survival Rate	100	100	100	98.5	96.7	94.5	92.0	89.1	86.0	82.7
Vintage	11	12	13	14	15	16	17	18	19	20
Survival Rate	79.1	75.4	71.6	67.7	63.7	59.7	55.7	51.8	47.9	44.2
Vintage	21	22	23	24	25	26	27	28	29	30
Survival Rate	40.6	37.1	33.7	30.6	27.6	24.8	22.2	19.8	17.6	15.5

The scrappage rates are taken from table 3.14 of the *Transportation Energy Data Book* (Davis et al., 2014) published by the Department of Energy. Vintage 0 trucks (not shown) survive with 100 percent probability.

C Details on the expected fuel savings calculations

In table C2 we report the survival rates used to compute lifetime fuel savings for trucks.

The scrappage rates are estimated using registration data on heavy trucks (i.e., trucks with a gross vehicle weight over 26,000 pounds), following the method described in Greenspan and Cohen (1996). We use the scrappage rates estimated for a 1980 model-year heavy truck.[38]

In table C3 we report nominal and real diesel prices. Nominal prices were downloaded from the Energy Information Administration and deflated using the personal consumption expenditure price index published by Bureau of Economic Analysis.

[38] In the same table are the estimated scrappage rates for a 1990 model year heavy truck. But the estimates for the 1980 model year seem more reasonable to us. This is because the median life of a 1990 model year heavy truck is estimated to be 28.0 years, a dramatic increase over the estimated life of 1970 and 1980 model year trucks, which are 20 and 18.5 years, respectively. Using the 1990 model year scrappage estimates would increase our estimated fuel savings from an increase in MPG, reinforcing our main result that trucks undervalue expected discounted savings from increased fuel efficiency.

Table C3: On-highway diesel fuel prices

Year	Nominal price	Real price
1992	1.084	1.084
1993	1.092	1.066
1994	1.101	1.052
1995	1.109	1.038
1996	1.235	1.132
1997	1.198	1.080
1998	1.044	0.934
1999	1.121	0.988
2000	1.491	1.282
2001	1.401	1.182
2002	1.319	1.098
2003	1.509	1.232
2004	1.810	1.443
2005	2.402	1.861
2006	2.705	2.042
2007	2.885	2.124
2008	3.803	2.717
2009	2.467	1.764
2010	2.992	2.104
2011	3.840	2.636
2012	3.968	2.675
2013	3.922	2.612
2014	3.825	2.514

Note: 1995 to 2015 nominal prices are published by the Energy Information Administration (EIA). The 1992 nominal price is derived from World Bank and EIA data on diesel prices. Specifically, the World Bank publishes U.S.diesel prices per liter, which we converted to gallons. The (converted) prices are \$1.060 and \$1.022 for 1992 and 1998 respectively. The EIA's published price in 1998 is \$1.044. Applying the percent change in World Bank prices to the EIA price, we arrive at a diesel price of \$1.084 in 1992. To arrive at 1993 and 1994 prices, we use a linear interpolation between 1992 and 1995. To arrive at real prices, nominal prices are deflated by the personal consumption expenditure price index, published by the Bureau of Economic Analysis.

D Estimation details

The solution to the weighted least squares problem (see equation 8) is given by

$$\hat{\omega}_j(p) = (X_D^T W X_D)^{-1} X_D^T W \left[\sum_{k=1}^{J} \mathbf{e}_k K_h(p_k - p)\right], \tag{D1}$$

where $\omega_j(p)$ is the vector $\big(\alpha_j(p), \lambda_j(p), \gamma_j(p)\big)$, $W = \text{diag}\{K_B(\mathbf{x_j} - \mathbf{x})\}$, e_k is the k^{th} unit vector, and X_D is the design matrix of equation (8) given by

$$X_D = \begin{pmatrix} 1 & (\mathbf{x}_1 - \mathbf{x})^T & \text{vech}((\mathbf{x}_1 - \mathbf{x})(\mathbf{x}_1 - \mathbf{x})^T)^T \\ 1 & (\mathbf{x}_2 - \mathbf{x})^T & \text{vech}((\mathbf{x}_2 - \mathbf{x})(\mathbf{x}_2 - \mathbf{x})^T)^T \\ \vdots & \vdots & \vdots \\ 1 & (\mathbf{x}_J - \mathbf{x})^T & \text{vech}((\mathbf{x}_J - \mathbf{x})(\mathbf{x}_J - \mathbf{x})^T)^T \end{pmatrix} \tag{D2}$$

In particular, the notation vech(A) denotes the half-vectorization of the symmetric matrix A. I.e., if A is an $l \times l$ matrix, then vech(A) is the vector in $\mathbb{R}^{l(l+1)/2}$ whose first l entries are the first column of A, whose subsequent $l-1$ entries are the second column of A below the diagonal, etc... From this solution, our estimate of the conditional density function follows immediately. It is just $\hat{g}_j(p|\mathbf{x}_j) = \alpha_j(p)$, and its derivatives are $\widehat{\frac{\partial}{\partial x_k} g_j}(p|\mathbf{x}_j) = \hat{\lambda}_{j,k}(p)$.

We can also note the expected value of the bracketed sum in equation (D1) is mechanically just the vector of observed prices. To see this, observe that the projection matrix $(X_D^T W X_D)^{-1} X_D^T W$ does not depend on the price. Moreover, since $K_h(p_k - p)$ is just a symmetric pdf centered at p_k, it must be the case that for any values of h and p_k, we have $E(K_h(p_k - p)) = p_k$. So, taking expected value of equation (D1),

$$
\begin{aligned}
E(\hat{\omega}_j(p)) &= (X_D^T W X_D)^{-1} X_D^T W \; E\left[\sum_{k=1}^{J} \mathbf{e}_k K_h(p_k - p)\right] \\
&= (X_D^T W X_D)^{-1} X_D^T W \left[\sum_{k=1}^{J} \mathbf{e}_k E\{K_h(p_k - p)\}\right] \\
&= (X_D^T W X_D)^{-1} X_D^T W \left[\sum_{k=1}^{J} \mathbf{e}_k p_k\right] \\
&= (X_D^T W X_D)^{-1} X_D^T W \mathbf{p}
\end{aligned}
$$

In other words, this conditional density estimator agrees with the more simple local quadratic estimator on the expected value of the price hedonic and its derivatives. However, the advantage of this more general framework is that it allows us to actually recover estimates of the unobserved characteristic. Obtaining estimates of the unobserved characteristic has proven essential when recovering the full utility function of each firm, including their willingness to pay for the discrete characteristics (Bajari and Benkard, 2005).

E Bandwidth selection and standard errors

In order to compute the optimal bandwidth h for the bandwidth matrix $B = h\mathbf{I}$, we use least squares cross-validation technique, as described in Fan and Gijbels (1996). Let $\hat{p}_h(\mathbf{x})$ denote the conditional estimate of the price hedonic at a point $\mathbf{x}$, using the bandwidth h. For each truck j, construct the leave-one-out estimate $\hat{p}_{h,-j}(\mathbf{x})$ by estimating the model on the subsample $\{p_i, \mathbf{x_i}\}_{i \neq j}$. We then examine and attempt to minimize the prediction error $p_j - \hat{p}_{h,-j}(\mathbf{x}_j)$. So we choose the bandwidth that minimizes the cross-validation function

$$
CV = \frac{1}{n} \sum_{j=1}^{n} \left[p_i - \hat{p}_{h,-j}(\mathbf{x}_j)\right]^2 w(\mathbf{x}_i), \qquad \text{(E3)}
$$

where the function $w(\cdot)$ is a weighting function that corresponds to the (scaled) inverse of the sample weights.

In our application, we select the optimal value of h numerically, by computing the cross-validation score for every bandwidth value between 0 and 10, at intervals of 0.01, and selecting the value of h which produces the minimum value. For our sample in 1992, this value is 2.48.

In order to compute the standard errors for the estimates of the distribution of taste parameters, we constructed bootstrap estimates by re-estimating the model on 10,000 random re-samplings of the data (sampled with replacement).

To compute the kernel smoothed distributions, we use N-fold (leave one out) cross-validation to select the bandwidth that minimizes the mean squared integrated error of our estimated distribution function. Formally, if $\hat{f}_h$ is the estimated distribution for a given bandwidth h, and data $\{X_i\}_{i=1}^n$, we partition the data into N equally sized sub-samples (consisting of a single observation) by assigning to each observation a sample $s(i)$. We then minimize the cross-validation function

$$CV(h) = \int \hat{f}_h^2(x)dx - \frac{2}{n}\sum_{i=1}^{n} \hat{f}_{h,-s(i)}(X_i), \tag{E4}$$

where $\hat{f}_{h,-s(i)}$ is the density function estimated using all of the data except the sub-sample to which observation i belongs.