Do truckers undervalue fuel efficiency? *

Jacob Adenbaum,† Adam Copeland,‡ and John Stevens§

April 17, 2015

Abstract

The U.S. federal government enacted fuel efficiency standards for heavy trucks for the first time in September 2011. The effectiveness of increasing fuel efficiency standards is an open question among economists; one rationale for using this policy tool is if there are informational frictions that result in vehicle purchasers undervaluing the savings from increased fuel efficiency (Parry, Walls, and Harrington 2007). We measure by how much heavy truck owners undervalue lifetime 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 heavy truck owners’ willingness to pay for fuel efficiency, with the elasticity of fuel efficiency to price ranging from 0.40 at the 10th percentile to 1.39 at the 90th percentile, and an average of 0.91. Combining these results with estimates of lifetime fuel savings from increases in fuel efficiency, we find that heavy truck owners’ willingness-to-pay for a 1 percent increase in fuel efficiency is, on average, just 21.7 percent of the expected lifetime fuel savings. These results suggest that introducing fuel efficiency standards for heavy trucks might be an effective policy tool.

Preliminary draft: Do not post, publish, or release in any form

*We thank Ed McCarty and John Miller for assistance with the data; Eugene Chung and Sarit Weisburd for research assistance. The views expressed here are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve Board, or the Federal Reserve System.

†Federal Reserve Bank of New York; e-mail: jacob.adenbaum@ny.frb.org
‡Federal Reserve Bank of New York; e-mail: adam.copeland@ny.frb.org (corresponding author)
§Federal Reserve Board; e-mail: john.j.stevens@frb.gov
1 Introduction

In September 2011 the U.S. federal government set fuel efficiency standards for medium and heavy trucks, for the first time. These regulations are a first step—the National Highway Transportation and Safety Administration (NHTSA) and the Environmental Protection Agency (EPA) are expected to announce a second round of fuel efficiency standards in March 2016. The motivation of these policy interventions is to reduce the impact of the negative externalities, such as air pollution, associated with using medium and heavy trucks, similar to the arguments made to set fuel efficiency standards for passenger cars (i.e. the Corporate Average Fuel Efficiency (CAFE) program).

The effectiveness of setting fuel efficiency standards to reduce these negative externalities is an open question (Parry, Walls, and Harrington 2007). One motivation for mandating fuel efficiency standards relies on the existence of informational frictions. Technology which increases fuel efficiency may not be adopted by manufacturers, for example, if consumers are not willing to pay for the increased cost of these technologies. This might be true even if the expected lifetime fuel savings from adopting the technology are greater than the adoption cost, because there are informational frictions which cause consumers to heavily discount expected fuel savings, such as short planning horizons. Greene (1998) posits that these informational frictions exist for the automobile and light truck market, and argues this is a reason why fuel efficiency standards (i.e., CAFE standards) have been effective in the U.S. In recent work, Allcott and Wozny (2014) present empirical evidence that consumers undervalue future gasoline costs.

As a first step toward understanding whether the new fuel efficiency standards on medium and heavy might be an effective policy instrument, we study whether truck owners undervalue the lifetime gains from increased fuel efficiency. We focus on class 8 trucks, a.k.a. highway rollers, which are the largest truck-tractors used on highways and 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. Consequently, a truck’s fuel efficiency should be reflected in its price. Furthermore, class 8 trucks are clearly defined and contain a

---

1See a summary released by the White House of this program along with other measures at [http://www.whitehouse.gov/sites/default/files/docs/finaltrucksreport.pdf](http://www.whitehouse.gov/sites/default/files/docs/finaltrucksreport.pdf).

2Need to cite various industry sources about the importance of fuel efficiency.
relatively small number of producers and products, especially compared with medium trucks. Lastly, there exist detailed data on the stock and value of used class 8 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 U.S. These data provide an extraordinary amount of information on trucks, because survey respondents, i.e. truck owners, answer a large number of questions on the physical and operational characteristics of their truck. In addition, this survey contains a host of demographic variables on each respondent, such as truck fleet size and major product hauled. Using truck characteristics, we are able to identify class 8 trucks in the survey. To assign prices to the highway rollers identified in the VIUS, we turn to the Truck Blue Book, a listing of truck prices conditional on characteristics. Based on truck characteristics, we are able to attach prices to trucks in the VIUS. The end result of this merger is a data set on the equilibrium price and quantity for a representative sample of highway rollers registered in the U.S. We have merged the VIUS and Truck Blue Book prices for the 1992 and 1997 census years. Our preliminary results focus on 1992.

We use these data to infer truck owners’ willingness-to-pay for truck characteristics, employing recent advances to the hedonic approach laid out in Bajari and Benkard (2005). Their results allow us to use the hedonic approach to recover willingness-to-pay in an environment with imperfect competition and the presence of a truck characteristic which is not observed by the econometrician; both of which are features of the heavy truck market. Their approach allows us to use a local quadratic method to infer truck owners’ willingness-to-pay for 4 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 for random coefficients, and so we estimate the distribution of willingness-to-pay for each continuous characteristic. Allowing for both heterogeneity in preferences and an unobserved characteristic are important aspects of our econometric approach; these features have been shown to

---

3 Prior to 1997 this survey was known as the Truck Inventory and Use Survey. The survey was discontinued after 2002.

4 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).
be crucial to arriving at accurate estimates of demand in a variety of applications (e.g., Berry, Levinsohn, and Pakes (1995) and Nevo (2001)).

Our estimates of the distribution of willingness-to-pay have the expected signs in that almost all truck owners positively value miles per gallon (MPG), engine size, and empty weight and negatively value lifetime miles. Focusing on MPG, 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 0.40 and 1.39 demonstrating that there is a wide range of tastes for fuel efficiency among truck owners.

To determine whether truck owners undervalue expected savings from fuel efficiency, we compute the discounted lifetime savings from a 1 percent increase in MPG. This measure varies across trucks, depending upon their current age and fuel efficiency. We then compare these estimated savings with owners’ willingness-to-pay and find that truck owners’ undervalue discounted lifetime savings from increased fuel efficiency. On average, owners are only willing to pay for 21.7 percent of discounted expected gains. The distribution of willingness-to-pay to lifetime fuel savings ranges from 4.7 percent at the 10th percentile to 42.4 percent at the 90th percentile. These results show that the vast majority of heavy truck owners undervalue expected lifetime fuel savings, which in turn suggests that imposing fuel efficiency standards might be an effective policy tool.

This paper builds upon a large literature focused on evaluating the effectiveness of fuel efficiency standards. As explained below in more detail, our contribution to this literature is twofold. First, we focus on class 8 trucks whereas previous work analyzes automobiles and light trucks. To our knowledge, this is the first paper to study whether heavy truck owners undervalue the expected discounted savings from increased fuel efficiency. Second, we use recent advances in the hedonic approach which allow us to identify the distribution of willingness-to-pay for observed characteristics using only cross-sectional data.

A number of papers have estimated whether households undervalue (or overvalue) the expected discounted gains from increased fuel efficiency. Goldberg (1998) finds that households do not undervalue the expected discounted fuel gains when considering purchases on new light vehicles, whereas more recent work by Allcott and Wozny (2014) finds that households do undervalue expected fuel savings.

\textsuperscript{5}Relatedly, Benkard and Bajari (2005) show that hedonic price indexes may be biased when not all product characteristics are observed.
dervalue expected savings. Our work differs in that we focus on heavy trucks as opposed to households purchasing automobiles and light trucks. Further, we implement a recent innovation to the hedonic approach. Following Bajari and Benkard (2005), we use local quadratic methods to estimate heavy truck owners’ willingness to pay for truck characteristics. This technique allows for random coefficients, providing us with an estimate of the distribution of willingness-to-pay for an characteristic (such as fuel efficiency). Furthermore, this method produces willingness-to-pay estimates even in the presence of an unobserved characteristic. This differs from past work, such as Espey and Nair (2005), which does not allow for random coefficients or account for an unobserved characteristic.

In the next section we introduce our model and then 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 highway rollers. A highway roller 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 4 continuous characteristics and dummy variables accounting for the truck manufacturer as well as the truck cab design. The continuous characteristics are miles-per-gallon (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 $x_j$ denotes a $1 \times K$ vector of physical characteristics, $p_j$ is the price of the truck, and $\xi_j$ is the unobserved characteristic. A truck owner $i$ maximizes utility by selecting a 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

\[
\max_{(j,c)} u_i(x_j, \xi_j, c)
\]

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 \( \tilde{p}(x, \xi_j) \) which maps the set of product characteristics to prices and satisfies \( p_j = \tilde{p}(x_j, \xi_j) \) for all \( j = 1, \ldots, J \) for a specific market and point in time. They also show that the unobserved characteristics can be identified using a single cross section: If \( (p, x, \xi) \) are distributed jointly with cumulative distribution function \( F(p, x, \xi) \), then the unobserved characteristics \( \xi_j \) are equal to the conditional cumulative distribution function of the prices,

\[
\xi_j = F_{p|x=x_j}(p_j).
\] (1)

To identify the price function and the unobserved characteristics, we assume \( \xi \) is independent of \( x \). An interpretation of this assumption is that the location of trucks in the characteristic space is exogenous, or at least determined prior to the revelation of consumers willingness-to-pay for \( \xi \). This is a reasonable assumption, because truck and engine re-designs, which involve substantial R&D, are often done every several years whereas \( \xi \) can vary on a more frequent basis. This assumption is slightly stronger than the mean independence assumption that is commonly used in the empirical industrial organization literature.

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

\[
u_{ij} = \beta_{ij}x_j + \beta_{i,\xi}\xi_j + c,\] (2)

where we use the log of the continuous characteristics: mpg, lifetime miles, engine size, and empty weight. This specification allows for random coefficients and notably does not impose a parametric assumption on the distribution of \( \beta_{ij} \) for the continuous characteristics.

---

\(^6\)The conditions are that (a) \( u_i \) is continuously differentiable in \( c \) and strictly increasing in \( c \) with \( \frac{du_i}{dc} > \epsilon \) for some \( \epsilon > 0 \) and all \( c \in (0, y_i) \); (b) \( u_i \) is Lipschitz continuous in \( (x_j, \xi_j) \); (c) \( u_i \) is strictly increasing in \( \xi_j \).

\(^7\)The unobserved characteristic may be influenced by marketing campaigns, for example.
A final assumption is that the choice set is continuous. Then, given an interior solution, the first order conditions imply

\[ \beta_{i,k} = x_{jk} \frac{d\tilde{p}}{dx_{jk}}, \quad (3) \]

where \( k \) indexes the four continuous characteristics. As detailed in section 4, the empirical challenge is to obtain estimates of \( \beta_{ik} \), the random coefficients, by estimating the derivative of the price function. According to our model, \( \beta_{ik} \) 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. The VIUS was a survey conducted every five years in order to track the stock of trucks operating in the U.S. (The survey was discontinued after 2002.) The Census surveyed a random sample of trucks registered or licensed in the U.S. 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), gas mileage, and lifetime miles. In addition, the VIUS records some information on the demographics of the owner of the truck. For example, information was recorded about the industry in which the truck was primarily used and the fleet size of the firm that owned the truck. This survey, then, provides a detailed look at the stock of trucks every five years as well as demographic information on the firms that own the trucks.

\(^{8}\)Before 1997, the VIUS was known as the Truck Inventory and Use Survey (TIUS). The Truck Blue Book is currently published by Primedia.
In this paper, we examine the 1992 and 1997 VIUS. Each survey is quite large; the 1992 and 1997 surveys contain 123,641 and 104,545 observations respectively. From these data on the stock of all trucks, we extracted the subset of trucks that fit our definition of highway rollers, 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 need to satisfy in order to be classified as a highway roller. 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 severely reduced the size of the VIUS in both census years, as observations on pickup trucks and other medium trucks (e.g., straight trucks or ‘box’ trucks) were eliminated. For the 1992 VIUS, for example, the sample size fell from 123,641 observations to 42,108. 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.\footnote{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.}

The first restriction mainly eliminated truck-tractors with two axles, as these trucks are limited by how much cargo they can pull. These trucks 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 an 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 is operated in a manner consistent with highway rollers, 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 for the 1992 VIUS and from 27,956 to 17,385 for the 1997 VIUS. We believe the resulting sample of observations is representative of the stock of highway rollers in the U.S. for the two census years.  

Although the VIUS provides a detailed accounting of the stock of highway rollers 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 truck prices based on their characteristics. From the publisher, we obtained the October issues of the *Truck Blue Book* for 1992 and 1997. We then assigned prices to trucks in the VIUS based on their recorded characteristics. In order to be able to compare prices in these two years, we deflated the 1997 price into 1992 dollars using the Bureau of Labor Statistics’ producer price index for heavy trucks.

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 these 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 VIN that reveals no confidential information from the Census for the 1992 and 1997 census years.

For 1992, the price data cover used trucks up to 8 years of age and for 1997 it covers trucks up to 9 years of age. This leaves us without prices for the oldest highway rollers. In 1992 and using the sample weights, these older trucks accounted for 29 percent of the stock of trucks. In 1997, the fraction is 22 percent. Due to the lack of price data, these older trucks are dropped from our analysis. To merge the newer trucks with prices, for each census year 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 matches of 75 and 76 percent in 1992 and 1997 respectively. For the trucks for which we could not find a

\footnote{A summary table of how we filtered the data is provided in the appendix.}
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. For both census years, 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 (8 and 9 percent of the trucks in 1992 and 1997, respectively). In the end, the final datasets used for analysis have 629 and 510 observations for the 1992 and 1997 census years, respectively.

### 3.2 Data Description

This work focuses on eight truck characteristics: year (i.e., 1992 or 1997), 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 miles per gallon (MPG). Using the VIUS, we identified the major brands (or makes) which produce trucks-tractors in the US. We further used industry information to consolidate brands which belonged to the same firm. In the end, we used 6 makes 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 US, conventional cabins are the dominant design. In the 1992 census year, conventional trucks make 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; conventional trucks make up 52 percent of trucks of vintage \(8^{[1]}\).

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 has a displacement of just under 212 cubic inches.) The empty weight of a truck is measured in pounds, lifetime mileage

---

\(^{[1]}\)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. In 1982, the federal government changed the regulation, and the length measurement focused on the trailer (and so excluded the truck-tractor).
Table 1: Truck Characteristics (1992)

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

Note: MPG is miles per gallon and SD is standard deviation. The mean and standard deviation statistics were computed using the 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.

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 characteristics, we compute a nearest neighbor statistic for each continuous characteristic. 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 edge of the product space which do not have close neighbors. For example, the three trucks with MPG of less than 4, and the handful with MPG above 8. But these products are a tiny part of the sample and so will not meaningfully impact 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 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 in the VIUS survey 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 in prices (as described in section 3.1). Although there are bunching of trucks in the highest engine size, figure 4 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 substantial variation in prices across trucks. Indeed, there is a large decline in prices across vintages; vintage 0 highway rollers have a median price of $48,875.
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.

and vintage 8 trucks have a price of $8,412 (see figure 5). Further, within each vintage there is a wide range of prices. The box-and-whiskers chart displayed in figure 5 illustrates the large inter-quartile price difference. For example, this difference 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 the price of truck (see table 2). For the 1992 census year, 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 by vintage increases as vintage decreases (i.e. as trucks get newer). The negative correlation
Note: A box-and-whiskers chart is displayed for each vintage in the 1992 census year. 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 continuous variables (1992)

<table>
<thead>
<tr>
<th></th>
<th>Price</th>
<th>MPG</th>
<th>Lifetime miles</th>
<th>Engine Size</th>
<th>Empty weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPG</td>
<td>0.17***</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifetime miles</td>
<td>−0.70***</td>
<td>−0.16***</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engine size</td>
<td>−0.08*</td>
<td>0.038</td>
<td>0.07</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Empty weight</td>
<td>0.11**</td>
<td>−0.13***</td>
<td>−0.005</td>
<td>0.003</td>
<td>1.00</td>
</tr>
</tbody>
</table>

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

between empty weight and MPG is likely driven by technology, in that making a truck lighter will increase its MPG.

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

\[
p(x_j, \xi_j) = \int p g(p|x_j) dp.
\]

Then taking advantage of linearity, we have

\[
\frac{\partial p_j(x_j, \xi_j)}{\partial x_{j,k}} = \frac{\partial}{\partial x_{j,k}} \int p \cdot g(p|x_j) dp, = \int p \cdot \frac{\partial g(p|x_j)}{\partial x_{j,k}} dp,
\]

where \( k \) denotes an element (a specific characteristic) in \( x_j \).

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, Yao, and Tong (1996). We can estimate \( g(p|x_j) \) using a nonparametric regression technique, because as \( h \to 0 \),

\[
E \left[ K_h(p_j - p)|x_j \right] \approx g(p|x_j),
\]

(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 around \( x_j \), we get

\[
E \left[ K_h(p_j - p)|x_0 \right] \approx g(p|x_j) + \frac{dg(p|x_j)}{dx_j} (x_0 - x_j) + \frac{1}{2} (x_0 - x_j)^T H(x_0 - x_j),
\]

(6)

where \( x_0 \) denote a vector of characteristics in the neighborhood of \( x_j \), \( H \) is the Hessian matrix of \( g(p|x_j) \) with respect to \( x_j \), and the superscript \( T \) denotes the transpose. Fan, Yao, and Tong (1996) redefine the right hand side as

\[
\alpha_j + \lambda_j^T (x_0 - x_j) + \gamma_j^T \text{vech} \left( (x_0 - x_j)(x_0 - x_j)^T \right),
\]

(7)

where \( \text{vech}(X) \) 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

\[
\min_{\alpha_j,\lambda_j,\gamma_j} \sum_{n=1}^J \left\{ K_h(p_n - p) - \alpha_j - \lambda_j^T (x_n - x_j) - \gamma_j^T \text{vech} \left( (x_0 - x_j)(x_0 - x_j)^T \right) \right\}^2 K_B(x_n - x_j),
\]

(8)

where \( K_B \) is a multivariate kernel weighting function with bandwidth matrix \( B \) such that \( K_B(u) = (1/|B|)K(||B^{-1}u||) \).\footnote{Note that if \( u \in \mathbb{R} \), then this reduces to the scaled kernel \( K_h \) since \( B \) has only one entry.} Our focus is on the price derivatives and so on the estimate of \( \lambda_j \). Nevertheless, we include the higher-order terms in the minimization problem to reduce the bias of the estimator.\footnote{See Racine (2008) for a primer on nonparametric estimation, including a discussion of bias and nonparametric estimators. See Fan and Gijbels (1996) for a detailed discussion of bias and local polynomial estimators.} Further, if we wanted to recover estimates of the unobservable characteristics \( \xi_j \) for each truck, we could do so by computing the sample analog to equation

14


or

$\xi_j = \int_{p < p_j} p \cdot g(p|\mathbf{x}_j) dp = \int_{p < p_j} p \cdot \alpha_j(p) dp.$  \hspace{1cm} (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 $B = hI$. 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. For this reason, we normalize each of our observed variables by its standard deviation for the purposes of computing the kernel weights (so that they are all on the same scale). 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 2.63. More details on least-squares cross-validation can be found in the appendix.

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 the appendix, 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.

4.2 Results

From our estimator, 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. These (kernel-smoothed) distributions are plotted in figures 6 to 9. Our willingness-to-pay
Table 3: Willingness-to-pay for truck characteristics (elasticities)

<table>
<thead>
<tr>
<th>truck characteristic</th>
<th>mean</th>
<th>standard deviation</th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miles per Gallon</td>
<td>0.91</td>
<td>0.43</td>
<td>0.40</td>
<td>0.66</td>
<td>0.92</td>
<td>1.16</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.17)</td>
<td>(0.25)</td>
<td>(0.24)</td>
<td>(0.23)</td>
<td>(0.22)</td>
<td>(0.22)</td>
</tr>
<tr>
<td>Lifetime Miles</td>
<td>-0.85</td>
<td>0.32</td>
<td>-1.21</td>
<td>-1.09</td>
<td>-0.92</td>
<td>-0.66</td>
<td>-0.37</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.18)</td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Engine Size</td>
<td>1.38</td>
<td>1.12</td>
<td>-0.02</td>
<td>0.73</td>
<td>1.47</td>
<td>2.10</td>
<td>2.69</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.11)</td>
<td>(0.26)</td>
<td>(0.38)</td>
<td>(0.50)</td>
<td>(0.59)</td>
<td>(0.65)</td>
</tr>
<tr>
<td>Empty Weight</td>
<td>0.45</td>
<td>0.27</td>
<td>0.17</td>
<td>0.31</td>
<td>0.46</td>
<td>0.62</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.21)</td>
<td>(0.15)</td>
<td>(0.15)</td>
<td>(0.15)</td>
<td>(0.15)</td>
<td>(0.15)</td>
</tr>
</tbody>
</table>

Note: MPG is miles per gallon. Standard errors were computed by bootstrapping.

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, 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.

The distribution of tastes for fuel efficiency is quite wide, with the 10th and 90th percentiles of the distribution equal to 0.40 and 1.39 respectively (see table 3). Owners also have a greater variety in tastes for lifetime mileage, the 10th and 90th percentiles are -1.21 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, 1.38 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 hand tail of the distribution hardly willing to pay for increased engine size, or even disvaluing it (at the 10th percentile, our elasticity estimate is -0.26) whereas those in the right tail have elasticities greater than 2 (at the 90th percentile, our elasticity estimate is
Figure 6: Distribution of willingness-to-pay for miles per gallon (1992)

Figure 7: Distribution of willingness-to-pay for lifetime miles (1992)
Figure 8: Distribution of willingness-to-pay for engine size (1992)

Figure 9: Distribution of willingness-to-pay for the weight of a truck when empty (1992)
In contrast, truck owners tastes for empty weight are more narrowly distributed, with an average elasticity of 0.45.

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 vintage, \( v_j \), and MPG, \( \text{mpg}_j \). Formally, these expected fuel savings, \( G_j \), are given by

\[
G_j = \sum_{v=v_j}^{L} \delta^{(v-v_j)} h(v_j, v) \left( \frac{m(v)}{\text{mpg}_j} \right) d(v-v_j)
\]

where \( L \) is the maximum possible lifetime of a truck and \( \delta \) is an annual discount rate of 0.95. The function \( h(v_j, v) \) is the probability that a truck of vintage \( v_j \) survives to age \( v \), \( m(v) \) is the expected mileage of a truck of vintage \( v \), and \( d(x) \) is the expected price of diesel \( x \) years in the future.

We use the Transportation Energy Data Book (Davis, Diegel, and Boundy 2014) for the survival rates of heavy trucks and, based on this schedule, set the maximum age of a truck, \( L \), to 30 years. We use the VIUS data to compute \( m(v) \), setting it equal to the average annual miles of a truck of vintage \( v \). Annual mileage declines with vintage, except when going from 0 to 1 vintage. This oddity is likely the result that truck owners of vintage 0 are more likely have driven their truck for less than a year, compared to owners of vintage 1, and so report a smaller annual mileage figure. Indeed, the average lifetime miles of vintage 0 trucks is greater than the average annual mileage. We nevertheless use average annual miles and so bias our measure of expected fuel savings for vintage 0 trucks downwards, which is the more conservative approach. Both the survival rates and average annual mileage are reported in the appendix, in table 7. Finally, we assume that truck owners view diesel prices as a random walk, and so expect future prices to be equal to the current price. Based on a combination of World Bank and EPA data on prices, we compute that the average price of diesel per gallon in
Table 4: Expected Lifetime Savings by Vintage and MPG (dollars)

<table>
<thead>
<tr>
<th>Vintage</th>
<th>5 MPG</th>
<th>6 MPG</th>
<th>7 MPG</th>
<th>8 MPG</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1,439</td>
<td>1,199</td>
<td>1,028</td>
<td>899</td>
</tr>
<tr>
<td>1</td>
<td>1,357</td>
<td>1,131</td>
<td>969</td>
<td>848</td>
</tr>
<tr>
<td>2</td>
<td>1,231</td>
<td>1,026</td>
<td>879</td>
<td>769</td>
</tr>
<tr>
<td>3</td>
<td>1,102</td>
<td>918</td>
<td>787</td>
<td>689</td>
</tr>
<tr>
<td>4</td>
<td>979</td>
<td>816</td>
<td>699</td>
<td>612</td>
</tr>
<tr>
<td>5</td>
<td>881</td>
<td>734</td>
<td>629</td>
<td>551</td>
</tr>
<tr>
<td>6</td>
<td>793</td>
<td>661</td>
<td>566</td>
<td>495</td>
</tr>
<tr>
<td>7</td>
<td>714</td>
<td>595</td>
<td>510</td>
<td>446</td>
</tr>
<tr>
<td>8</td>
<td>639</td>
<td>532</td>
<td>456</td>
<td>399</td>
</tr>
</tbody>
</table>

Note: MPG is miles per gallon. Each cell reports the expected lifetime savings of a 1 percent increase in MPG, conditional on a truck’s vintage and current MPG.

From these data, we compute the expected fuel savings corresponding to a 1 percent increase in fuel efficiency. To provide a sense of the magnitude of these savings, in table 4 we report our calculations of the expected lifetime savings for specific vintage and mileage pairs.

Using our estimates of the price elasticities with respect to fuel efficiency, we calculate each truck owner’s willingness to pay for such an increase. We then take the ratio of a truck owner’s willingness to pay over the expected lifetime fuel savings, to measure by how much truck owners undervalue expected 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 10.

As illustrated in figure 10, truck owners heavily undervalue lifetime fuel efficiency savings. On average, owners are willing to pay for only 21.7 percent of the lifetime fuel cost savings that accrue from a 1 percent increase in MPG. At the 90th percentile the distribution, owners are willing to pay for 42.4 percent of these lifetime fuel savings whereas at the 10th percentile owners are willing to pay for only 4.7 percent the savings (see table 5). It is only at the

---

15The 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 EPA’s published price in 1998 is $1.045. Applying the percent change in World Bank prices to the EPA price, we arrive at a diesel price of $1.084 in 1992.
Figure 10: Distribution of truck owners’ valuation of lifetime fuel savings (1992)
rightmost tail of the distribution do we find owners willing to pay for the full lifetime fuel savings.

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 lifetime fuel savings than owners of older trucks. We plot the cdf of the distribution of owner’s valuation of lifetime fuel savings conditional on the owner’s truck being of vintage 0 (figure 11) and vintage 8 (figure 12). In comparing these two cdfs, we see that the median vintage 0 truck owner is willing to pay for roughly 50 percent of lifetime fuel savings, whereas the median vintage 8 owner is only willing to pay about 10 percent.

6 Conclusion

In this paper, we estimate that truck owners of highway rollers 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 0.24 (10th percentile) to 1.40 (90th percentile). On average, trucks owners are willing to pay 0.83 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 potential gains from increased fuel efficiency. On average, we find that owners are willing to pay for only 20.0 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 might be an effective policy tool.
Figure 11: Cdf of vintage 0 truck owners’ valuation of lifetime fuel savings

Figure 12: Cdf of vintage 8 truck owners’ valuation of lifetime fuel savings
References


24


Table 6: Filtering of the 1992 and 1997 census year data

<table>
<thead>
<tr>
<th></th>
<th>observations</th>
<th>sample weights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nmbr</td>
<td>sum</td>
</tr>
<tr>
<td>1992 census year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIUS data set</td>
<td>123,641</td>
<td>5,920,075,519</td>
</tr>
<tr>
<td>Only truck-tractors</td>
<td>42,108</td>
<td>116,817,124</td>
</tr>
<tr>
<td>Have three axles</td>
<td>32,240</td>
<td>83,909,899</td>
</tr>
<tr>
<td>Have a conventional or cabin-over-engine design</td>
<td>31,335</td>
<td>81,696,386</td>
</tr>
<tr>
<td>Have diesel engine and air brakes</td>
<td>30,591</td>
<td>79,208,583</td>
</tr>
<tr>
<td>Do not spend most of their time off-road</td>
<td>29,588</td>
<td>77,244,933</td>
</tr>
<tr>
<td>Has the correct body type</td>
<td>26,668</td>
<td>70,086,661</td>
</tr>
<tr>
<td>1997 census year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIUS data set</td>
<td>104,545</td>
<td>72,800,251,891</td>
</tr>
<tr>
<td>Only truck-tractors</td>
<td>27,956</td>
<td>1,543,752,184</td>
</tr>
<tr>
<td>Have three axles</td>
<td>21,749</td>
<td>1,140,559,205</td>
</tr>
<tr>
<td>Have a conventional or cabin-over-engine design</td>
<td>20,611</td>
<td>1,074,940,667</td>
</tr>
<tr>
<td>Have diesel engine and air brakes</td>
<td>19,867</td>
<td>1,030,056,985</td>
</tr>
<tr>
<td>Do not spend most of their time off-road</td>
<td>19,292</td>
<td>1,005,571,131</td>
</tr>
<tr>
<td>Has the correct body type</td>
<td>17,385</td>
<td>919,831,799</td>
</tr>
</tbody>
</table>

Note: nmbr is number and pct is percent.

7 Appendix

7.1 Data details

In this section we provide details on how the data used for analysis was constructed. In table 6, we report how the different filters we used to identify highway rollers in the VIUS reduced the size of the sample, both in terms of the number of observations and the sum of the sample weights.
7.2 Expected fuel savings calculations

In Table 7 we report the expected miles and survival rates used to compute lifetime fuel savings for trucks. Expected miles are taken from the 1992 VIUS, and are the average annual miles reported by respondents for the subset of trucks classified as highway rollers. The 1992 VIUS records a truck’s model year for vintages 0 through 9. Vintage 10 is a catch-all category for all trucks older than vintage 9. In the VIUS, we compute that trucks of vintage greater than 9 have an average annual mileage of 35,065. Because we focus on trucks of early vintages, for our fuel savings calculations only the average annual mileage for vintages 10 through 28 is important. As such, we do not try to interpolate a (more reasonable) decline in annual miles for these older trucks.

The scrappage rates are taken from Table 3.14 of the Transportation Energy Data Book (Davis, Diegel, and Boundy 2014) published by the Department of Energy. The scrappage rates are estimated using registration data on heavy trucks (i.e., trucks with a gross vehicle weight over 26,000 pounds), using the method described in Greenspan and Cohen 1996. We use the scrappage rates estimated for a 1980 model year heavy truck.

7.3 Estimation details

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

$$\hat{\alpha}_j(p) = (X_D^T W X_D)^{-1} X_D^T W \left[ \sum_{k=1}^{J} e_k K_h(p_k - p) \right], \quad(11)$$

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

---

16In 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 7: Expected Miles and Survival Rates by Vintage

<table>
<thead>
<tr>
<th>Vintage</th>
<th>Expected Annual Mileage (miles)</th>
<th>Survival Rate (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>82,145</td>
<td>100.0</td>
</tr>
<tr>
<td>1</td>
<td>102,567</td>
<td>100.0</td>
</tr>
<tr>
<td>2</td>
<td>100,823</td>
<td>100.0</td>
</tr>
<tr>
<td>3</td>
<td>94,021</td>
<td>100.0</td>
</tr>
<tr>
<td>4</td>
<td>85,839</td>
<td>98.5</td>
</tr>
<tr>
<td>5</td>
<td>79,397</td>
<td>96.7</td>
</tr>
<tr>
<td>6</td>
<td>73,020</td>
<td>94.5</td>
</tr>
<tr>
<td>7</td>
<td>69,541</td>
<td>92.0</td>
</tr>
<tr>
<td>8</td>
<td>63,351</td>
<td>89.1</td>
</tr>
<tr>
<td>9</td>
<td>56,179</td>
<td>86.0</td>
</tr>
<tr>
<td>10</td>
<td>35,065</td>
<td>82.7</td>
</tr>
<tr>
<td>11</td>
<td>35,065</td>
<td>79.1</td>
</tr>
<tr>
<td>12</td>
<td>35,065</td>
<td>75.4</td>
</tr>
<tr>
<td>13</td>
<td>35,065</td>
<td>71.6</td>
</tr>
<tr>
<td>14</td>
<td>35,065</td>
<td>67.7</td>
</tr>
<tr>
<td>15</td>
<td>35,065</td>
<td>63.7</td>
</tr>
<tr>
<td>16</td>
<td>35,065</td>
<td>59.7</td>
</tr>
<tr>
<td>17</td>
<td>35,065</td>
<td>55.7</td>
</tr>
<tr>
<td>18</td>
<td>35,065</td>
<td>51.8</td>
</tr>
<tr>
<td>19</td>
<td>35,065</td>
<td>47.9</td>
</tr>
<tr>
<td>20</td>
<td>35,065</td>
<td>44.2</td>
</tr>
<tr>
<td>21</td>
<td>35,065</td>
<td>40.6</td>
</tr>
<tr>
<td>22</td>
<td>35,065</td>
<td>37.1</td>
</tr>
<tr>
<td>23</td>
<td>35,065</td>
<td>33.7</td>
</tr>
<tr>
<td>24</td>
<td>35,065</td>
<td>30.6</td>
</tr>
<tr>
<td>25</td>
<td>35,065</td>
<td>27.6</td>
</tr>
<tr>
<td>26</td>
<td>35,065</td>
<td>24.8</td>
</tr>
<tr>
<td>27</td>
<td>35,065</td>
<td>22.2</td>
</tr>
<tr>
<td>28</td>
<td>35,065</td>
<td>19.8</td>
</tr>
<tr>
<td>29</td>
<td>35,065</td>
<td>17.6</td>
</tr>
<tr>
<td>30</td>
<td>35,065</td>
<td>15.5</td>
</tr>
</tbody>
</table>
\[
X_D = \begin{pmatrix}
1 & (x_1 - x)^T & \text{vech}((x_1 - x)^T(x_1 - x)^T) \\
1 & (x_2 - x)^T & \text{vech}((x_2 - x)^T(x_2 - x)^T) \\
& \vdots & \vdots \\
1 & (x_n - x)^T & \text{vech}((x_n - x)^T(x_n - x)^T)
\end{pmatrix}
\]

(12)

In particular, the notation \(\text{vech}(A)\) denotes the half-vectorization of the symmetric matrix \(A\). I.e., if \(A\) is an \(l \times l\) matrix, then \(\text{vech}(A)\) 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, the our estimate of the conditional density function follows immediately. It is just \(\hat{g}(p|x_j) = \alpha_{j,0}(p)\), and its derivatives are \(\frac{\partial}{\partial x_{k}} \hat{g}(p|x_j) = \hat{\lambda}_{j,k}(p)\).

We can also note the expected value of the bracketed sum in equation (11) is mechanically just the vector of observed prices. To see this, observe that the projection matrix \((X_D^TWX_D)^{-1}X_D^TW\) 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 (11),

\[
E(\alpha_j(p)) = (X_D^TWX_D)^{-1}X_D^TW \left[ \sum_{k=1}^{J} e_k K_h(p_k - p) \right]
\]

\[
= (X_D^TWX_D)^{-1}X_D^TW \left[ \sum_{k=1}^{J} e_k E(K_h(p_k - p)) \right]
\]

\[
= (X_D^TWX_D)^{-1}X_D^TW \left[ \sum_{k=1}^{J} e_k p_k \right]
\]

\[
= (X_D^TWX_D)^{-1}X_D^TW \bar{p}
\]

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. See Bajari and Benkard (2005).
7.4 Bandwidth Selection and Standard Errors

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

$$CV = \frac{1}{n} \sum_{j=1}^{n} \left[ p_i - \hat{p}_{h,-j}(x_j) \right]^2 w(x_i),$$

(13)

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).