Gasoline Prices and the Fuel Economy Discount Puzzle

Hunt Allcott and Nathan Wozny

September 18, 2009

Abstract

It is often asserted that consumers purchasing automobiles and other durable goods underweight gasoline or other future add-on costs. We test this hypothesis in the US automobile market by examining the equilibrium effects of time series variation in gasoline price expectations on the market shares and relative prices of vehicles with different fuel economy ratings. When gas prices rise, gas guzzlers become less desirable while higher fuel economy vehicles become more attractive. We should therefore expect the relative price of the former to go up. However, market share changes - increased production of high fuel economy vehicles and scrappage of low fuel economy vehicles - can attenuate this effect. We identify the causal impact of gas prices on willingness-to-pay for fuel efficiency using a discrete choice framework that captures substitution patterns between new and used vehicles. This framework suggests a natural instrument for the market share of a used vehicle: the interaction of fuel economy with gasoline prices in the year the vehicle was produced. We construct an unprecedentedly rich dataset that includes the prices, quantities, and characteristics of all new and used vehicles on the road monthly from 2003 to 2008. We show that the American auto consumer is willing to pay just $0.25 to reduce future gas expenditures by $1 (in present discounted value). Our finding provides support for a paternalistic justification of interventions such as energy efficiency standards: they force consumers to purchase energy-efficient goods, regardless of whether their demand patterns indicate that they want them.

We thank Alberto Abadie, Orley Ashenfelter, Lucas Davis, Kelly Gallagher, Penny Goldberg, Michael Greenstone, Mark Jacobsen, James Kahn, Lutz Kilian, David Laibson, Henry Lee, Robin Lee, Lars Lefgren, Alessandro Lizzeri, Erich Muehlegger, Sendhil Mullainathan, Richard Newell, Ariel Pakes, Jesse Rothstein, Jim Sallee, Rob Stavins, Chris Timmins, Roger von Haefen, Matt White, and seminar participants at Harvard, MIT, Princeton, and Stanford for their feedback. Funding for this project was provided by Princeton’s Industrial Relations Section, Harvard’s Energy Technology Innovation Policy group and the Harvard University Center for the Environment. We thank Carmen Collyns and Jordon Ricks for research assistance on this and related projects. We especially thank Souman Hong for meticulous and tireless research assistance.
1 Introduction

There is a growing body of evidence that consumers choosing between products may underweight, relative to purchase prices, product costs that are less salient or accrue in the future. Consumers on eBay, for example, are less elastic to shipping and handling charges than to the listed purchase price (Hossain and Morgan 2006). Mutual fund investors appear to be less sensitive to ongoing management fees than to upfront payments (Barber, Odean, and Zheng forthcoming). Consumers seem to discount highly the future energy costs of air conditioners relative to their purchase prices (Hausman 1979).

Similarly, it is often asserted that vehicles’ gasoline costs are not salient to automobile consumers at the time of purchase, and that they thus do not fully consider these costs when choosing between automobiles. As a result, "myopic" consumers lock themselves into lower fuel economy automobiles, with higher resulting fuel expenditures, than they would optimally choose. In 2007, the median-income American family spent $2400 on gasoline, and American households spent $286 billion in total (U.S. Bureau of Labor Statistics 2007). Misoptimization over such a large expenditure class could result in substantial welfare losses. Such myopia would also help explain what Jaffe and Stavins (1994) call the "Energy Paradox": that consumers and firms have been remarkably slow to adopt energy efficient technologies.

The welfare losses from myopia would be exacerbated by externalities related to national security and climate change. There has been substantial debate over whether these should be internalized through gasoline taxes or Corporate Average Fuel Economy (CAFE) standards (e.g. Bento, et al, 2009). Economists typically argue that gas taxes are preferable because they act both on the extensive margin, by encouraging consumers to buy higher fuel economy vehicles, and on the intensive margin, by encouraging them to drive vehicles less. Although CAFE standards only bind on the extensive margin, the myopia argument suggests that consumers’ intensive margin response to gasoline taxes is not optimal. CAFE standards therefore could increase welfare by

---

1 See, for example, Greene (1998) and Parry, Walls, and Harrington (2007).
2 Various explanations have been proposed for this apparent anomaly, including imperfect information (Metcalf 1998), credit constraints (Pizer, et al, 2002), principal-agent problems (Murtishaw and Sathaye 2006), bounded rationality (DeCanio 1993), and that discount rates do not properly model hysteresis and irreversible investment under uncertainty (Metcalf and Rosenthal 1995).
3 The higher fuel economy vehicles required under CAFE require less fuel to operate per mile, and thus consumers actually have the perverse incentive on the intensive margin to increase use. This is often called the "rebound effect."
forcing consumers to own higher fuel economy vehicles, regardless of whether they think they want them.

A central problem in this discussion is the dearth of evidence on whether automobile consumers actually are or are not myopic. This paper uses unprecedentedly comprehensive micro and market-level data on the prices, quantities, characteristics, and usage of vehicles in the United States to ask, *Are automobile consumers myopic with respect to future gasoline costs?* A risk neutral neoclassical consumer should be indifferent between a one dollar decrease in the vehicle purchase price and a one dollar decrease in the expected present value of fuel costs. "Myopia" has taken on several meanings; for our exposition, we define myopia as the revealed willingness to trade off a dollar in discounted future gasoline costs for less than a dollar in purchase price.

Our empirical work is based on the idea that time series changes in expected gasoline prices should differentially affect demand for vehicles with different fuel economy ratings, and thus their equilibrium prices and quantities. Using panel data on vehicle markets with vehicle-specific fixed effects allows our estimator to be consistent even if a vehicle’s fuel economy is correlated with other observed and unobserved vehicle characteristics. The panel approach is simplest when we assume that neither new vehicle supply nor used vehicle scrappage rates respond to gas prices, as in Kahn (1986). In this model, a vehicle’s price relative to any other vehicle should decrease by one dollar for each one-dollar increase in the relative present discounted value of expected future gasoline costs.

However, the expected response of vehicle supply to gas price changes biases that potential approach towards detecting myopia. If higher gas costs lead to a decrease in production of low fuel economy new vehicles, the supply shift would have a positive effect on prices of these new vehicles, partially offsetting the effect of decreased demand. New vehicle prices would therefore change less than expected if quantities were assumed to be fixed. A similar effect on used vehicle

---

4 This "paternalistic" argument for fuel economy standards is made by Greene (1998), Fischer, Harrington, and Parry (2007), Parry, Walls, and Harrington (2007), among others.

5 In Gabaix and Laibson (2006), for example, a myopic consumer is an individual who does not fully analyze the future game tree. In game theory, it typically means agents that place zero weight on future periods. In other settings, myopes have short time horizons and/or high discount rates.

prices would occur if scrappage of low fuel economy vehicles increased with gas prices.\(^7\) A further concern is that a supply response in the new vehicle market affects the prices of used vehicles that are substitutes: a decrease in production of new gas guzzlers will likely increase demand for similar used gas guzzlers, again offsetting the direct effect of a gas price increase on used vehicle prices.

We address the need to account for vehicle quantities and substitution patterns using a discrete choice model of vehicle demand that allows for unobserved heterogeneity in consumers’ preferences. Consumers’ utility from owning a vehicle is a function of both discounted future gasoline costs and the purchase price. Aggregated to the market level, these demand functions give a relationship between equilibrium vehicle prices, market shares, and gasoline costs. Because vehicle quantities (and prices) are correlated with unobserved vehicle characteristics, we instrument for equilibrium quantities by exploiting the fact that the demand for a new vehicle with low fuel economy is lower in years when gasoline prices are high. At any time in the future, the quantity available of the (now used) vehicle produced in that year will therefore be lower than the quantity of the same model sold in a year when gasoline prices were low. Crucially, this within-model variation in quantity of used vehicles should be uncorrelated with unobserved product attributes. We can therefore introduce a new instrument for the quantity of a used vehicle on the road: the interaction of fuel economy with the gasoline price in the year in which the vehicle was new. For new vehicles, we instrument with vehicle attributes using the approach of Berry, Levinsohn, and Pakes (1995).

We estimate the model with an unusually rich dataset on vehicle prices, quantities, and attributes between 2003 and 2008. From microdata on 55 million vehicle transactions at both auto dealerships and auctions, we construct monthly average prices for all new and used passenger vehicles available in the United States. From national-level data on vehicle registrations, we observe the market shares of each of these vehicles and match these to the price data using the industry’s serial numbers, called VINs. This is in turn matched to each vehicle’s fuel economy and other characteristics. The vehicle-level data are supplemented by data on retail gasoline prices and oil futures prices, from which we construct expected future gasoline costs, and large nationally-representative surveys of vehicle ownership and vehicle-miles traveled.

\(^7\)Indeed, media reports and academic analyses have documented that as gasoline prices rose between 2003 and 2008, the market shares of new high fuel economy vehicles rose, (Klier and Linn 2008) the scrappage of used low fuel economy vehicles increased (Li, Timmins, and Von Haefen 2009), and the relative prices of both new and used vehicles with low fuel economy dropped (Busse, Knittel, and Zettelmeyer 2009).
Over our sample period, we find that consumers are indifferent between a reduction of automobile purchase price by one dollar and a reduction of discounted future gasoline costs by four dollars. The order of magnitude of this finding is consistent across time periods and if choice sets are restricted to different subsets of vehicle ages and classes. This finding is robust also to different assumptions on vehicle survival probabilities, usage, and other parameters.

These findings have implications in several domains. The apparent myopic behavior towards gasoline costs suggests that a gasoline tax is not sufficient to correct negative externalities associated with gasoline consumption. Understanding consumers’ demand for fuel economy is central to analyzing the welfare and profit implications of new products and regulatory changes in the automotive industry. Literature including Allcott and Muehlegger (2008), Bento, et al (forthcoming), Berry, Levinsohn, and Pakes (1995, 2004), Goldberg (1998), Jacobsen (2008), and Nevo (2002) all depend on an estimate of consumers’ demand for fuel economy. Our analysis is complementary to this body of work in that it provides a careful estimate of an essential demand parameter.

Evidence that consumers are myopic with respect to future product costs also has important implications for how firms behave in equilibrium. Myopia is one reason that can explain why firms set low markups on base products such as credit card interest rates, razors, and printers and high markups for add-ons such as late fees, blades, and ink cartridges (Gabaix and Laibson 2006). In the auto industry, although manufacturing firms are not involved in the "aftermarket" for gasoline, the fuel economy embodied in their products determines gasoline demand, and increasing vehicle fuel economy is costly. Consumer myopia in this domain, if it indeed exists, reduces manufacturers’ ability to exploit economies of scale in producing high fuel economy vehicles and dulls their incentives to invest research and development funds in reducing the cost of these vehicles. This suggests additional channels through which regulations such as Corporate Average Fuel Economy standards and improved fuel economy labels at automotive dealerships can affect consumer welfare.

The paper progresses as follows. In section 2, we provide an overview of how we think about this problem, including both econometric identification issues and conceptual modeling challenges. In the third section, we formally set up consumers’ utility functions, and in Section 4, we present our estimation strategy. Section 5 presents the aggregate and consumer-level data at our disposal. In section 6, we discuss the construction of discounted expected gasoline costs, including different
formulations of consumers’ discount rates and expectations over future gasoline prices. Section 7 presents our results, and section 8 concludes.

2 An Overview of the Problem

All else equal, a consumer should be willing to pay $1 more for a vehicle that costs $1 less to operate over the vehicle’s lifetime. The fundamental goal of this paper is to test whether observed automobile market equilibria are consistent with this consumer behavior. Our test requires us to construct a framework that predicts how gasoline price-induced demand shifts affect equilibrium vehicle prices, and then to compare the predicted vehicle prices to data. This section presents a simple, intuitive framework for doing this. Our formal discrete choice model will be presented in the following section.

A number of analyses have attempted to test for myopia in the purchase of energy using durable goods - or similarly to compute the "implicit discount rate" for future energy costs such that consumers appear trade these off equally with purchase prices - by exploiting variation in the prices and energy efficiencies of a cross section of choices. Cross-sectional identification strategies have been employed in a discrete choice framework in a seminal paper by Hausman (1979) and a number of later papers, including Espey and Nair (2005), Dreyfus and Viscusi (1995), and Dubin (1992). In the simplest version of this approach, consumers’ willingness to pay for fuel efficiency and other characteristics is estimated in a hedonic regression.

For such an estimator to be unbiased, any unobserved characteristics must be uncorrelated with energy cost, and the functional form of any observed and correlated characteristics must be correctly specified. With automobiles, this is likely to be problematic. Fuel economy is strongly correlated with weight and horsepower, which enter the typical indirect utility function for automobiles in characteristics space. While these variables are observable, the way in which they enter the utility function could be mis-specified. Furthermore, fuel economy is affected by styling decisions that

---

8 A second way of inferring discount rates is engineering estimates of the installation and operating costs of energy efficiency investments, as in Anderson and Newell (2004). In automobiles, this would require comparing the marginal cost of fuel efficiency upgrades that manufacturers choose to perform to the marginal fuel savings.

9 At least since Atkinson and Halvorsen (1984), it has been pointed out that the high correlation between weight and fuel economy, makes it difficult to estimate demand for fuel economy. In fact, cross sectional specifications of automobile demand in characteristics space sometimes give the wrong sign on fuel economy.
affect wind resistance and may enter utility functions, as well as by features such as air conditioners that increase a vehicle’s value. These features are in some cases difficult to quantify for the purposes of including them in a utility function. In a cross section, fuel economy is negatively correlated with price, which suggests that low fuel economy vehicles may have more unobserved characteristics that enter utility functions.

The ability to look "within" the same vehicle over time as gasoline prices change exogenously obviates the need to assume anything about how the vehicle’s characteristics are correlated with expected future gasoline costs. Even with panel data, the effect of changing gasoline costs must be separated from the effect of the vehicle’s age and any market-level time trends. We avoid making assumptions about depreciation patterns by comparing vehicles of the same model and age across different years with different gasoline prices. For example, compare the price of a model year 2001 Honda Civic in 2006 to a model year 2002 Honda Civic in 2007. Holding all else equal other than gasoline prices, the change in price of this 5 year old Civic from 2006 to 2007 is the response to the change in expected gas costs over that year.¹⁰

What follows in this section is not our formal model, but putting some mathematical structure on the problem can begin to build intuition. Consider a world with a fixed supply of a single vehicle with no substitutes. A simple demand function for the vehicle is given by:

\[ q = \alpha_0 - \eta P \]  \hspace{1cm} (1)

In this equation, \( P \) is the total operating cost of the vehicle. This cost can be written as \( P = p + G \), where \( p \) is the purchase price and \( G \) is the gasoline costs over the vehicle’s lifetime, which we assume to be the same for all consumers. The variable \( G \) will depend on future gas prices, fuel economy, and the vehicle’s usage and scrappage probability over time, and we will later return to this issue in great detail. Suppose we allow consumers to value purchase price and gas costs unequally:

¹⁰An alternative to exploiting the panel data is to compare prices within a model and model year, which requires imposing structure on depreciation patterns. Kahn (1986) follows this approach.
\[ q = \alpha_0 - \eta p - \gamma G \]

We can also re-arrange this equation to get equilibrium price on the left-hand side:

\[ p = -\frac{\gamma}{\eta} G + \frac{1}{\eta} (\alpha_0 - q) \]

Since supply is fixed, if \( \eta = \gamma \), then a one dollar increase in \( G \) causes a one dollar decrease in \( p \). More realistically, suppose that there are two substitutable vehicles, but vehicle 2 has lower fuel economy and therefore has higher gasoline costs. A simple demand equation for vehicle 2 relative to vehicle 1 (i.e. the number of consumers who would be willing to pay a given premium \( p_2 - p_1 \) to switch to vehicle 2 from vehicle 1) is:

\[ q_2 = \alpha_0 - \eta(p_2 - p_1) - \gamma(G_2 - G_1) \]

Rearranging this gives the relative price of vehicle 2 compared to vehicle 1:

\[ p_2 - p_1 = -\frac{\gamma}{\eta}(G_2 - G_1) + \frac{1}{\eta} (\alpha_0 - q_2) \]

This above equation shows that if \( \gamma = \eta \), a one dollar relative increase in \( G_2 \) causes a one dollar relative decrease in \( p_2 \). More generally, for a fixed supply of many vehicles:

\[ p_j = -\frac{\gamma}{\eta} G_j + \psi_j \]

In this equation, \( \psi_j \) is a constant for each vehicle \( j \), equal to a marginal consumer’s willingness to pay to operate the vehicle. If we observe the same set of vehicles in the same fixed supply over time as gas prices change, then \( \frac{\gamma}{\eta} \) can be identified by a panel regression:
\[ p_{jat} = -\frac{\gamma}{\eta} G_{jat} + \psi_{ja} + \xi_{jat} \]  \hspace{1cm} (7)

Since vehicles depreciate as gas costs change, we observe the “same” vehicle over time by comparing vehicles of the same model and same age, hence the additional \( a \) subscript and idiosyncratic variation across model years \( \xi_{jat} \). Performing this OLS regression with used vehicle data, described in detail in section 5, we estimate \( \frac{\gamma}{\eta} = 0.08 \), suggesting that prices move substantially less than expected. This is qualitatively consistent with the analogous specifications in Kahn (1986) and the literature that follows his approach.\(^{11}\)

Now suppose that consumers have demand for a single vehicle given by (2), but the supply of that vehicle is elastic:

\[ q = \beta_0 + \beta_1 p \] \hspace{1cm} (8)

Then the equilibrium vehicle price is given by:

\[ p = -\frac{\gamma}{\eta + \beta_1} G + \frac{\alpha_0 - \beta_0}{\eta + \beta_1} \] \hspace{1cm} (9)

Since upward sloping supply gives \( \beta_1 > 0 \), a one dollar increase in gas costs results in a less than one dollar decrease in vehicle price even when \( \gamma = \eta \). Thus, estimating equation (7) when supply is in fact elastic would result in a downward bias in an estimate of \( \frac{\gamma}{\eta} \), making possible a false rejection of the hypothesis that vehicle prices fully adjust to gas price changes. One response is to analyze only used vehicles, assuming the effect of gas price expectations on scrappage rates is negligible. However, similar new and used vehicles are likely to be substitutes. If an increase in gas prices leads to a decrease in production of new low fuel economy, this will increase the willingness to pay for a used low fuel economy vehicle that is a good substitute. This results in a positive

\(^{11}\)Working papers by Kilian and Sims (2006) and Sallee and West (2008) also report qualitatively similar results. After running the analogous specification, Kahn (1986) further experiments with different formulations of how consumers update gasoline price expectations, and eventually concludes that vehicle prices fully adjust to gasoline price changes.
correlation between $G_{jat}$ and $\xi_{jat}$ in equation (7), leading again to a downward bias in an estimate of $\frac{\gamma}{\eta}$.

In order to test for myopia in light of the confounding effects of supply shifts, we must build a model that allows quantities to vary endogenously, and also captures substitution patterns across similar vehicles. Our model characterizes the demand side of market equilibrium,\footnote{Our analysis uses only one of the necessary conditions for equilibrium provided by the demand side of the market. Although the supply side of the new vehicle market should also respond in equilibrium to gasoline price changes, a model of new vehicle suppliers in equilibrium is not needed to test for myopia on the demand side.} which generates predictions about how equilibrium vehicle prices and quantities respond to changes in gasoline prices. We can then test these predictions using market data.

3 Model

In this section, we describe our logit model, which is a modification of the standard framework in the industrial organization discrete choice literature, e.g. Berry (1994). In our static discrete choice model, consumers $i = 1, ..., N$ derive utility from owning and driving vehicles and from consumption of an outside good. In each period, indexed by $t$, consumers have homogeneous expectations $E[g_{t+1}, g_{t+2}, ..., \Omega_t]$ about the future path of gasoline prices $g$ given information set $\Omega$. We assume that consumers are risk-neutral. Given that the variance in total gasoline expenditures driven by uncertain gasoline prices, while a non-trivial amount of money, is still typically small relative to the auto owner’s overall wealth, we think that this is a reasonable assumption.

Consumers choose from a set of new and used models $j = 1, ..., J_t$, each with characteristics $x_{ja}$, where $a$ indexes the vehicle’s age. Consumers also can choose an outside option, denoted $j = 0$, which is to own no vehicle and instead walk or take public transit. Consumer $i$ receives utility $u_{ijat}$ from purchasing vehicle $ja$ in year $t$:

$$u_{ijat} = -\eta p_{jat} - \gamma G_{jat} + \psi_{jat} + \epsilon_{ijat}$$

(10)

where $p_{jat}$ is the purchase price and $G_{jat}$ is the discounted present value of total future gasoline costs. If consumers value purchase price and gas costs equally, then $\eta = \gamma$. $G_{jat}$ depends on the
discount rate, expected future gasoline prices, and expected usage of the vehicle; we will return to this quantity in more detail in section 5. \( \psi_{jat} \) is the present discounted value of the flow utility that vehicle \( ja \) will provide to the average consumer over the rest of its lifetime from year \( t \) forward.

Given the importance of substitution across vehicle models and ages to our analysis, it is crucial that we correctly model the relationship between a consumer’s taste for a particular vehicle and her taste for another similar model. A consumer who has a particular taste for one vehicle, for example, may also have a taste for vehicles in the same class (such as mid-size sedans). This motivates our use of a nested logit model, which allows correlations in the utilities of vehicles within predefined groups. As in a simple logit model, we assume that \( \epsilon_{ijat} \) is distributed Type I Extreme Value. This produces a tractable relationship between market-level prices and shares. To capture substitution patterns, we allow a consumer’s idiosyncratic preferences to be correlated across vehicles within the same predetermined group: \( \text{corr}(\epsilon_{ijat}, \epsilon_{ij'a't}) \) is nonnegative when \( ja \) and \( j'a' \) are in the same group and zero otherwise. We will estimate a parameter \( \sigma \) related to these within-group correlations\(^{13}\).

As in other nested logit models, the nests are specified ex ante and determine the structure of substitution patterns allowed by the model. Vehicles are divided into disjoint sets comprised of vehicles over which the analyst believes are close substitutes. This division may occur along multiple dimensions, such as vehicle class or age. We use class as the first nest because this substitution is central to our analysis: consumers are unlikely to have equal preferences for vehicles of substantially different sizes, and failing to account for this substitution pattern would likely lead us to overstate consumers’ responsiveness to gas price changes, biasing the test in favor of detecting myopic behavior\(^{14}\).

Normalizing the utility of the outside good to zero and aggregating the nested logit across

\(^{13}\)In particular, the joint distribution of \( \epsilon_{ijat} \) for all \( ja \) for individual \( i \) at time \( t \) is: \( f(\cdot) = \exp \left[ - \sum_{k \in K} \left( \sum_{jw \in B_k} e^{-(\epsilon_{ijw't}/(1-\sigma))} \right)^{1-\sigma} \right] \). \( K \) is the set of all groups of vehicles, and \( B_k \) is the set of vehicles in group \( k \). \( \sigma \) is a parameter related to the within-group correlation of utilities and will be estimated in the model. As \( \sigma \) approaches one, the within-group correlation of utilities approaches one. If \( \sigma = 0 \), the standard logit model is recovered. This distribution can be extended to multiple nests.

\(^{14}\)The difference between this model and a random coefficients model is that the latter allows unobserved taste shocks associated with continuous measures of product characteristics, whereas the former allows unobserved taste shocks only over dummy variables indicating membership in pre-defined groups. In some sense, the random coefficients models are more flexible, but in principle both capture different forms of unobserved heterogeneity. The econometric appeal of the nested logit model is that it allows market level prices and quantities to be described as an additive log-linear function, which will be crucial to our instrumentation and estimation strategy. The benefits of this approach will become clear in the "Empirical Strategy" section.
consumers gives a well-known equation for market shares (e.g. Berry 1994):

\[
\ln s_{jat} - \ln s_{0t} = -\eta p_{jat} - \gamma G_{jat} + \sigma \ln s_{(ja/k)t} + \psi_{jat}
\]  

(11)

In this equation, \( s_{jat} \) is the market share of vehicle \( ja \) out of all vehicles, \( s_{0t} = \) is the share of the outside option, and \( s_{(ja/k)t} \) is the share of \( ja \) out of all vehicles in group \( k \).

The market-level relationship between equilibrium prices and quantities implied by this discrete choice framework will be the basis of our empirical test of myopia. Specifically, we wish to test for whether the marginal utility of the net present value of money for gasoline purchases equals the marginal utility of money for the vehicle purchase:

\[
\text{Neoclassical Consumers : } H_0 : \frac{\gamma}{\eta} = 1
\]

\[
\text{Myopic Consumers : } H_A : \frac{\gamma}{\eta} < 1
\]  

(12) (13)

We assume that \( \eta \) and \( \gamma \) are homogeneous in the population, which produces this simple hypothesis that can be tested with a linear model. In reality, consumers have different marginal utility of money, which explains differences in preferences for luxury vehicles and vehicles of different ages. By including nests for age classes and luxury vehicles, however, our nested logit specification does allow unobserved differences in utility for vehicles in different price ranges. In an empirical extension, we also allow \( \frac{\gamma}{\eta} \) to vary depending on vehicle MPG.

Appendix 9.1 provides formal detail on the assumptions required to derive this static model from a dynamic consumer choice model. Once the model and estimation strategy have both been outlined, we will return to these and other assumptions and clarify their costs and benefits.

---

15While we do not allow heterogeneity in \( \eta \) and \( \gamma \), we note that many discrete choice models do allow for heterogeneous marginal utility of money. The appeal of the typical approach is discussed in Berry, Levinsohn, and Pakes (1995): it captures the fact that consumers have heterogeneous wealth and declining marginal utility of money and predicts supply side markups that are more realistic because they increase in price. While using the nested logit error structure to proxy for this is somewhat ad-hoc, we think it is sensible in our application, both because we need not simulate markups and because we benefit substantially from being able to set up a simple linear test.

12
4 Empirical Strategy

The typical discrete choice estimation would proceed by estimating equation (11) with market share as the dependent variable, using some instrument for price to control for the correlation with unobserved characteristics. If $\hat{\eta} > \hat{\gamma}$, we would reject $H_0$. Instrumenting for price, however, increases the standard error on $\hat{\eta}$, reducing the power of this test.

A critical step in our empirical strategy follows from the observation that while textbook derivations of aggregate market shares from individual choice probabilities typically conclude with market shares on the left-hand side, this need not be the actual estimating equation. The equation (11) is simply a relationship between prices and shares that is necessary for vehicle demand to be in equilibrium. We can re-arrange this equation by moving shares to the right hand side, moving price to the left hand side, and dividing by $\eta$:

$$p_{jat} = -\frac{1}{\eta} (\ln s_{jat} - \ln s_{0t}) - \frac{\gamma}{\eta} G_{jat} + \frac{\sigma}{\eta} \ln s_{(ja/k)t} + \frac{\psi_{jat}}{\eta}$$ (14)

The first term is likely to be endogenous, as market share should be correlated with unobserved quality, and we will momentarily introduce an appropriate instrument. Importantly, however, this equation gives a new test of myopia, whether $\frac{\hat{\gamma}}{\hat{\eta}} < 1$. As we will discuss, we need not instrument for $G$, which reduces the standard error on $\frac{\hat{\gamma}}{\hat{\eta}}$ and increases the power of our test.

In general, average utility obtained from a vehicle $\psi_{jat}$ depends on average preferences for observed and unobserved characteristics. It is theoretically possible to estimate equation (14) using a cross section of vehicles with different prices and fuel economy ratings. This requires, however, that we can observe and parameterize vehicle characteristics well enough that no unobserved part of $\psi_{jat}$ is correlated with fuel economy. Our panel identification strategy allows us to relax the assumption that vehicle fuel economy is uncorrelated with the econometric error by exploiting model-by-age fixed effects $\psi_{ja}$ rather than including observable characteristics. Since observable vehicle characteristics are, by definition, effectively identical across the years of vehicle $ja$, the deviation from vehicle average utility is a year-specific unobservable: $\psi_{jat} = \psi_{ja} + \xi_{jat}$. Equation (14) becomes:
\[ p_{jat} = -\frac{1}{\eta} (\ln s_{jat} - \ln s_{0t}) - \frac{\gamma}{\eta} G_{jat} + \frac{\sigma}{\eta} \ln s_{(ja/k)t} + \frac{\psi_{ja}}{\eta} + \frac{\xi_{jat}}{\eta} \]  

(15)

The fixed effect allows us to relax the assumption that \( G_{jat} \) is uncorrelated with observed and unobserved product characteristics that are constant across model years, i.e. to allow that \( E[G_{jat}] \neq 0 \). Unbiasedness still requires that within-\( ja \) deviations \( \Delta G_{jat} \) are uncorrelated with unobserved characteristics, i.e. that \( E[\Delta G\xi] = 0 \).

Equation (15) resembles the reduced form equation (7) from Section 2. Were we willing to assume that market shares are fixed, or more weakly, uncorrelated with \( G \), then \( \frac{\gamma}{\eta} \) could be identified as the coefficient of \( G_{jat} \) in an OLS estimation of equation (7). This is a clear problem, as there is substantial evidence that both new vehicle sales and used vehicle scrappage respond to gasoline prices, and our model needs to account for this. Recall from above, however, that to estimate equation (15), we need an instrument that generates variation in market shares that is uncorrelated with quality. Interestingly, the fact that new vehicle sales respond to gasoline prices suggests an instrument for used vehicle quantities. Put differently, our problem is also part of the solution.

### 4.1 Instruments and Two-Stage Least Squares

Our instrument for quantities exploits the stylized fact that initially motivated the discrete choice model: vehicle market shares respond to gasoline prices. In particular, in years when gasoline prices are high, more high fuel economy vehicles are sold. This difference in quantities in use then persists over time: the increase in gasoline prices from 2004 to 2005, for example, means that there should be more two-year old gas guzzlers on the road in 2006 compared to 2007. Crucially, this effect is independent of the unobserved characteristic \( \xi_{jat} \).

Our instrument for the market shares of used vehicles is the expected lifetime gasoline costs of model \( j \) in year \( t-a \), when the vehicle was new, denoted \( G_{j0(t-a)} \). For within-nest market share, one intuitively would like to know how the expected lifetime fuel cost when vehicle \( j \) was new compares to the average for vehicles in nest \( k \), which we denote \( G_{k0(t-a)} \). If \( G_{j0(t-a)} - G_{k0(t-a)} > 0 \), then vehicle \( j \) had above average expected fuel costs; the more positive that number, the lower its within-nest market share should be. Since the term \( G_{j0(t-a)} \) is already included as our first instrument,
our additional instrument for within-nest market share is $G_{k0(t-a)}$. Since some production decisions may be made in advance, we also include as instruments expected gas costs (and their within-nest mean) in year $t - a - 1$. For expositional clarity, we omit these in the text below.

These instruments are not valid for new vehicles, as the proposed instrument is also the right hand side variable $G_{jat}$. Instead, we use as instruments for new vehicles the vehicle’s characteristics, the sum of these characteristics over all other vehicles produced by the same manufacturer, and the sum of these characteristics over all vehicles produced by competing manufacturers. This follows Berry, Levinsohn, and Pakes (1995), except for a Cournot instead of Bertrand model of differentiated product oligopoly\(^{16}\).

The new vehicle instruments are denoted by $z_{j0t}$. The new vehicle instruments are multiplied by a new vehicle dummy, and the gas price instruments are multiplied by a used vehicle dummy. Including month-by-year time dummies $\tau_t$ to soak up market level shocks, the first stage equations of the two stage least squares regression is:

\[
(\ln s_{jat} - \ln s_{0t}) = (\alpha_{11} G_{j0(t-a)} + \alpha_{12} G_{k0(t-a)}) \cdot 1(a > 0) \\
+ \alpha_{13} z_{j0t} \cdot 1(a = 0) + \alpha_{14} G_{jat} + \psi_{ja} + \tau_t + \varepsilon_{jat} \\
\ln s_{(ja/k)t} = (\alpha_{21} G_{j0(t-a)} + \alpha_{22} G_{k0(t-a)}) \cdot 1(a > 0) \\
+ \alpha_{23} z_{j0t} \cdot 1(a = 0) + \alpha_{24} G_{jat} + \psi_{ja} + \tau_t + \varepsilon_{jat} \tag{16}
\]

The second stage is:

\[
p_{jat} = \frac{1}{\eta} (\ln s_{jat} - \ln s_{0t}) + \frac{\sigma}{\eta} \ln s_{(ja/k)t} - \frac{\gamma}{\eta} G_{jat} + \frac{\psi_{ja}}{\eta} + \tau_t + \frac{\xi_{jat}}{\eta} \tag{17}
\]

\(^{16}\)To use the Berry, Levinsohn, and Pakes (1995) instruments, however, we need to assume that $E[z\xi] = 0$. In words, this is that the unobserved attribute is uncorrelated with product characteristics such as weight, which may themselves be correlated with $G$. We recognize the problem that this is inconsistent with our earlier identification assumptions that motivated the panel approach with fixed effects.
4.2 Revisiting Modeling Assumptions

Recall that our identification strategy requires that $E[\Delta G\xi] = 0$, where $\xi_{jat}$ is the deviation in year $t$ from vehicle $ja$’s average utility $\psi_{ja}$. Having completed the description of our estimator, we can now revisit the modeling assumptions and ask whether they are likely to bias our results. In each case, we will either conclude that the assumption is "conservative," in the sense of biasing our results against concluding that consumers are myopic, or discuss an empirical test that rules it out.

Taken literally, our logit model implies that a consumer makes a static decision to purchase a vehicle with the intention of holding it indefinitely. In reality, however, the consumer may sell her vehicle in any period at the market price and purchase a different vehicle. In Appendix 9.1, we show the steps necessary to derive our static discrete choice utility function from a dynamic model of consumer behavior.

Some may find it counterintuitive to compute $G_{jat}$ based on the expected fuel costs over the life of the vehicle, whereas in reality most consumers will eventually resell the vehicle and will therefore not incur all of these herself. The derivation in Appendix 9.1 makes clear why a consumer should care about lifetime fuel costs: when she resells the vehicle, the resale price should also account for the fuel costs over the remaining vehicle life.

The critical assumption required in that derivation is that consumers believe the current vehicle market to be "stationary." While this can be weakened slightly, our analysis does require that consumers expect changes in future market shares to be uncorrelated with past changes in gas price expectations. This assumption allows us to substantially simplify how consumers will form expectations of a vehicle’s future resale price. Market shares do, however, respond to gas prices, and if gas prices rose between the previous period and this period, sales of new high fuel economy vehicles increase. As these vehicles age, the supply of high fuel economy vehicles at any given age will now be higher, and their prices lower. Because this attenuates the increase in prices of high fuel economy that we should observe when gasoline prices increase, assuming it away biases the estimation in favor of detecting myopia.

Both of these concerns center on how consumers formulate the expected resale value of a vehicle. In our empirical analysis, however, we will show that issues related to resale values cannot account for the full scope of the mispricing that we observe: even when we make the alternative assumption
that consumers consider $G$ only over the time horizon of the typical vehicle holding period, we still find $\frac{\hat{\sigma}}{\sigma} < 1$.

While the purchase of a vehicle is not irreversible, in a dynamic model, transactions costs make the decision costly to reverse. An increase in uncertainty should cause consumers to delay purchasing vehicles, reducing vehicle demand and thus prices. Because the change in variance of $G_{Jat}$ induced by an increase in the variance of gasoline prices is higher for low fuel economy vehicles, this effect should be stronger for gas guzzlers. Risk aversion, instead of the risk neutrality that we assume, should magnify this effect\textsuperscript{17}. Since increases in gasoline prices over the past few years were also associated with increases in implied volatility, uncertainty and risk aversion should reduce the relative prices of low fuel economy vehicles as gasoline prices increased. Since we find that relative prices of low fuel economy vehicles did not reduce as much as our model predicts, the puzzle remains even in the face of both uncertainty and risk aversion.

For consistency, we need that consumers’ preferences are constant, or more weakly, that changes in preference parameters $\psi$ are uncorrelated with changes in gasoline prices. This assumption is fundamental to our identification strategy, which uses utility fixed effects for each vehicle over time, and requires that deviations from the utility fixed effect are uncorrelated with changes in the cost of fueling the vehicle. The assumption would be violated if consumers grew more fond of high fuel economy vehicles over the study period. This may indeed have occurred: an increase in environmentalism, for example, could have increased demand for hybrids as a signal of environmentalist identity. Importantly, however, this example biases us away from concluding that consumers are myopic: it is an unaccounted effect that should increase the relative price of high fuel economy vehicles, while we find that the relative price of high fuel economy vehicles did not increase as much as the model predicts.

\textsuperscript{17}In our static logit model, risk aversion also biases us in a conservative direction: increases in the variance of the distribution of $G$ that accompanied higher gasoline price expectations reduce the appeal of low fuel economy vehicles. This unaccounted effect should reduce their relative prices. Since we find that the observed relative prices of low MPG vehicles did not drop as much as our model predicts, the puzzle remains.
5 Data

We have assembled from multiple sources a comprehensive data set of the average prices, quantities, and characteristics of all passenger vehicle models registered in the US, in monthly cross sections from January 2003 to December 2008. Our dataset comprises 620,000 observations, which can be divided into 25,000 model-by-age groups. Table 1 presents descriptive statistics, and the Data Appendix (section 9) provides extensive detail on variable construction.

Used vehicle prices are based on auction data obtained from Manheim, the largest automobile auctioneer in the United States. We have data on each of the approximately 5 million vehicle sales that occur annually through Manheim auctions, which accounts for half of the country’s auction volume. We use the individual auction data to predict the mean price of each model in each month, adjusting for the vehicle's condition, odometer reading, and region and method of sale. New vehicle prices are from the Power Information Network, a network of dealerships managed by JD Power and Associates. These dealerships report 2.5 million new vehicle transactions each year, about 15% of the nation’s market. For each model, we observe monthly mean prices adjusted for consumer cash rebates and the difference between the negotiated trade-in price and the trade-in vehicle’s actual resale value, if any. Mean new and used vehicle prices for selected model years are shown in Figure 10.1. Although new vehicle prices are substantially higher than prices of used vehicles sold early in the vehicle’s life, this discontinuity will not affect our analysis.

We observe national-level quantities in use of each vehicle model on an annual basis from July 2003 through July 2008. These data are from the National Vehicle Population Profile, which we purchased from the automotive market research firm R.L. Polk. The quantities represent all vehicles registered as of July 1 of each year, including individual owners and fleets such as taxis, rental cars, and corporate and government motor pools. A vehicle may be driven on public roads only if it is registered, so this database is exhaustive for all intents and purposes. Total registered quantities for selected model years are shown in Figure 10.2.

Fuel economy data are from the U.S. Environmental Protection Agency (EPA), which has estimated the miles per gallon of all new vehicles since 1974. The EPA uses a test to determine fuel economy over a standardized drive cycle and then adjusts the results to account for the typical
consumer’s in-use fuel economy. The distribution of fuel economy (in miles per gallon) is shown for selected model years in Figure 10.3. Vehicle classes, which are used to define nests in the nested logit model, are also taken from the EPA’s fuel economy dataset. All other vehicle characteristics are from the Ward’s Automotive Yearbook.

Data on new and used vehicle prices and registered quantities are matched using digits from the Vehicle Identification Number that are common to all vehicles within a model and model year. Fuel economy data and other characteristics are matched by the vehicle make, model year, model name, trim level, body style, and the engine’s displacement and number of cylinders.

To estimate Vehicle-Miles Traveled (VMT), we use publicly-available data in the National Household Travel Survey for 2001 (and 2008, when available in late 2009). These are nationally-representative surveys of approximately 25,000 households that report, among many other variables, the age, fuel economy, and vehicle class for each of the household’s vehicles. As part of the survey, about 25,000 vehicles in the national sample had their odometers read twice, with several months in between readings. Figure 10.4 illustrates these data by showing the average annualized value of measured VMT by vehicle age.

For spot gasoline prices, we use monthly data on U.S. City Average Motor Gasoline Retail Prices from the US Energy Information Administration. As will be discussed in the following section, we also use Light Sweet Crude Oil futures prices from the Intercontinental Exchange (ICE) and New York Mercantile Exchange (NYMEX) to construct expectations of future gasoline prices.

6 Expected Discounted Gasoline Costs

The previously-introduced variable $G_{jat}$ is the net present value of expected lifetime gasoline costs over future years $s$:
\[ G_{jat} = E_t \left[ \sum_{s=t+1}^{t+(L-1-a)} \beta^{s-t} \cdot g_s \cdot m_{ja} \cdot f_{jas} \cdot \phi_{jas} \right] \]

\[ = \sum_{s=t+1}^{t+(L-1-a)} \beta^{s-t} \cdot E_t[g_s] \cdot E_t[m_{ja}] \cdot E_t[f_{jas}^{-1}] \cdot E_t[\phi_{jas}] \]  \hspace{1cm} (18)

\( L \) denotes the maximum lifetime of a vehicle, which we take to be 25 years. \( g_s \) is a gasoline price in year \( s \), \( m_{ja} \) is expected vehicle miles traveled, \( f_{jas} \) is fuel economy in miles per gallon, \( \phi_{jas} \) is the probability that the vehicle survives to year \( s \) conditional on surviving to its current age,\(^{18}\) and \( \beta \) is an annual discount factor. We model that all gasoline costs flow at the end of the period. This increases the time discounting for each flow, thus decreasing \( G_{jat} \) and the changes therein, which is conservative in the sense that it biases us away from finding myopia.

The second line of equation (18) includes expectations of separate quantities - gasoline prices, fuel economy, VMT, etc. - that we will derive from separate datasets. To get the product of these separate expectations in the second line from the expectation of products in the first line requires that these variables are uncorrelated. This is untrue for vehicle miles traveled \( m \) and gasoline prices \( g \): in reality, consumers are price elastic on the intensive margin, and there is a large literature documenting how vehicle use responds to gasoline prices (e.g. Hughes, Knittel, and Sperling 2007). If owners of low fuel economy vehicles respond to high gas prices by driving less, then we overstate their gas costs. However, the consumers’ utility from vehicle use also decreases in response to the reduced driving. The Envelope Theorem implies that these changes are equal and opposite to first order (Kahn 1986), and in our primary specifications, we assume away these effects. At the end of this section, however, we also derive an alternative specification that directly accounts for the effects of intensive margin elasticity.

Equation (18) admits heterogeneity in \( G_{jat} \) across vehicles, but not across consumers within vehicles. In reality, differences across consumers in the proportion of city versus highway driving, how fast they drive and accelerate, and how well they maintain their vehicles generate differences in

\(^{18}\) We assume that vehicles survive with probability one throughout each period, then a fraction determined by \( \phi_{jas} \) are removed from the market at the end of the period.
realized fuel economy. Furthermore, there is in reality substantial variation in vehicle-miles traveled across consumers that own the same vehicle.

We abstract away from this within-jat heterogeneity in $G$ for two reasons. First, it allows us to estimate market-level regressions, i.e. with observations at the jat level, instead of individual-level moments. Second, it is not obvious how in theory it should affect the results. An increase in gasoline prices may cause a consumer who drives an above-average distance each year to have a higher than average demand for a high fuel economy vehicle. Similarly, a consumer with a below-average VMT will find low-MPG vehicles relatively more attractive compared to the average consumer. Intuitively, we would expect these demand shifts approximately to cancel. Whether this heterogeneity across consumers causes a bias in our model’s predicted prices depends on the distribution of VMT and its correlation with other preferences for vehicle characteristics. However, we are not aware of a plausible distribution that causes such a concern.\footnote{If this "reshuffling" of vehicles across consumers is empirically relevant, we should see an increase in the correlation of individual-level VMT and fuel economy as gas prices increase. We will test this by examining individual-level vehicle ownership and use data from the 2001 and 2008 National Household Travel Surveys, once the latter becomes available later this year.}

In this section, we describe the formulation of the different components of $G_{jat}$. In sensitivity analyses to presented later, we will show that our qualitative conclusion is unaffected by plausible alternative assumptions. Readers interested in even more detail on the construction of $G$ may also consult the Data Appendix.

### 6.1 VMT and Survival Probability

Vehicle-miles traveled and expected lifetime vary significantly across vehicles, over time, and across consumers of the same vehicle. Importantly, the vehicle-level expectations of these variables may be correlated with the vehicle’s fuel economy. For example, American-made cars tend to be less durable than foreign cars, and they also have lower fuel economy. Accounting for this correlation is required to avoid bias, and closer approximations of the vehicle-level averages of these two variables improves efficiency.

To predict VMT for all vehicles in our sample, we take the annualized odometer readings from the 2001 NHTS and regress them on age, class, and fuel economy. We then take these estimated
coefficients and fit a predicted VMT to all the model-by-model year-by age observations required for the computation of $G_{jat}$. Note again that this produces a VMT conditional on 2001 gasoline prices; in an alternative specification to be described later we adjust VMT (and vehicle utility) for the gasoline price at time $t$.

We compute survival probability based on observed survival probabilities in our registration data. As with VMT, we assume that these survival probabilities does not depend on expected gasoline prices. Since our market data on prices controls for the vehicle’s odometer reading, we need not worry about changes in the vehicle’s expected life due to gasoline price induced shifts in driving behavior.

### 6.2 Discount Rates

Equation (18) assumes a particular functional form by which consumers discount future years of expected gasoline costs. In principle, this discounting may take on any functional form, and our model could be extended to allow for this. In practice, as discussed in the results section, automobile market data do not appear to have the power needed to reliably estimate this function. However, choosing an appropriate discount rate is essential to ensuring that any conclusions we draw from estimates of $\frac{\gamma}{\eta}$ are accurate. In particular, assuming a lower discount rate than consumers actually face would inflate $G_{jat}$, biasing our tests towards detecting undersensitivity to gas costs. Put differently, an implicit discount rate could be computed by adjusting $r$ until $\frac{\gamma}{\eta} = 1$.

In reality, consumers have a range of discount rates $r = \beta^{-1} - 1$ that they should apply to future flows of gasoline costs. In this section, we propose two ways to pin down what these discount rates should be. For background, data we acquired from an industry research firm (Spinella 2008) indicates that approximately half of new and used vehicles are bought on credit, while half are bought with cash.

The first approach is to view the purchase of a higher fuel economy vehicle as an investment in gasoline: in states of the world with high gasoline prices, the consumer has more money than in the counterfactual in which she had bought a gas guzzler. The Capital Asset Pricing Model (CAPM) is the most common way of determining the risk-adjusted rate of return that the consumer would require for such an investment. In this model, consumers choose a portfolio of investments that
will maximize expected utility given declining marginal utility of money. Declining marginal utility implies that assets with high expected returns in states of the world when the market is performing well are not as valuable in utility terms as assets with high expected returns when the market performs badly.

Annual data from 1919 to 2003 show that oil prices (and therefore gasoline prices) are slightly negatively correlated with market returns, as measured by the S&P 500 stock index.\(^{20}\) This negative correlation makes a high fuel economy vehicle an attractive investment, because the consumer pays less for gasoline when her other investments are down and marginal utility of money is high. Indeed, the Capital Asset Pricing Model predicts that the consumer should demand a rate of return on this investment that is less than the risk-free rate of return. Measured by the real return to US Treasuries, this has typically been close to one percent. By this approach, the risk-adjusted discount rate should be close to zero, and \(\beta\) therefore close to one. We note that although we assumed earlier that consumers are risk neutral, using CAPM to compute risk-adjusted interest rates is one approach to incorporating risk-aversion into the analysis.

A second approach to determining an appropriate discount rate is to look at consumers’ marginal cost of capital from loans. For consumers buying autos on credit, this would likely be the interest rate on their auto loan. The real interest rate on auto loans tends to be 12-15 percent (see, e.g. Dreyfus and Viscusi (1995), U.S. Federal Reserve Board (2007)). To conservatively bias ourselves against finding that \(\frac{\gamma}{\eta} < 1\), we take the upper end of reasonable discount rates: we assume a homogeneous discount rate of 15 percent and perform sensitivity analyses with 5 and 25 percent.

\(^{20}\) The formula we use for this calculation is:

\[
E[R_{Gas,t}] = R_{f,t} + \beta_{Gas} \cdot (E[R_{Market}] - R_{f,t})
\]

In this equation, \(R_{Gas,t}\) is the real rate of return to the "investment in gasoline" in year \(t\), \(R_{f,t}\) is the risk-free rate in year \(t\), \(\beta_{Gas} = \frac{cov(R_{Gas,t}, R_f)}{var(R_{Market})}\), and \(R_{Market}\) is the market rate of return. Using annual data from 1919 to 2003, we compute that \(\beta_{Gas} = -0.04\). Cochrane (2001) provides an intuitive explanation of CAPM and is an excellent source of further detail.
6.3 Gasoline Price Expectations

Just as there is likely to be some discount rate for which consumers appear to account for expected gasoline costs correctly, there are infinitely many forms of future gasoline price expectations for which vehicle price changes are consistent with rational, forward-looking consumers. However, finding that \( \hat{\beta} = 1 \) for an implausible set of expectations of future gasoline prices certainly is not evidence that consumers account for future gas costs correctly. While we cannot know with certainty what consumers believe about gasoline prices, we can test whether \( \hat{\beta} = 1 \) for a set of plausible expectations of future gasoline prices.

Gasoline prices move very closely with crude oil prices: Light Sweet Crude Oil spot prices predict 93% of the monthly variance in gasoline prices. This implies that crude oil futures prices are very good proxies for the market’s expectation of future gasoline prices.\(^{21}\) This subsection describes a procedure to use oil futures prices as a proxy for gasoline price expectations.

Oil futures prices are transformed into gasoline price expectations in 2005 dollars using the following approach. First, all data are deflated into real July 2005 dollars. Second, since purchasing a futures contract entitles the owner to receive a future amount of money at the settlement date, the futures price must be further deflated to account for inflation expectations between the trade date and the settlement date. Inflation expectations are backed out from the relative prices of inflation-protected versus standard Treasury bonds. Finally, the expected crude oil spot price in 2005 dollars is transformed to a gasoline price using the relationship between crude oil and gasoline spot prices fitted in a linear regression with an intercept. Table 10.5 shows the annual average gasoline spot prices and crude oil futures prices transformed to dollars per gallon of gasoline.

Although these contract are only traded with high liquidity with settlement dates less than two to three years in the future, the table illustrates that there are some trades observed for settlement dates as far as ten years in the future. This still does not allow a full characterization of gasoline price expectations over the life of the vehicle: there is some probability of survival for much longer than ten years, although gasoline costs will be substantially discounted in these out years at the

\(^{21}\)It has been suggested to us that the path of oil futures may not be a good measure of the path of consumers’ gasoline price expectations. It seems implausible that the futures market prices for a commodity would differ substantially from the price expectations of the principal users of that commodity. If our results are interpreted as finding either that consumers are myopic or that the auto market’s expectations of oil prices differ substantially from the oil market’s expectations, we believe that this is still a remarkable anomaly.
discount rate we use. As illustrated in Figure 10.5, as gas prices rose between 2003 and 2008 above their 1990’s average of approximately $1.50 per gallon, the futures market expected prices to eventually return closer to that previous level. We exploit that fact by postulating a simple model of mean reverting expectations where deviations at time $t$ from the $1.50$/gallon mean decay exponentially over years $s$:

$$E[g_{t+s}] = 1.50 + (g_t - 1.50) \cdot e^{\rho s}$$

Re-arranging this equation, the mean reversion parameter $\rho$ can be estimated from the post-1991 observed futures data using the following linear specification:

$$\log |Eg_s - 1.50| = \log |g_t - 150| + \rho(t - s)$$

The equation fits the data very well: it explains 85% of the variance in the observed futures data between 1999 and 2008. The estimation gives $\hat{\rho} = -0.057$, meaning that the market expected recent gasoline price increases to decay back to $1.50$ per gallon at 5.7 percent per year.

One important takeaway is that the market does not believe that gasoline prices are a martingale: while we could perhaps estimate some other form of expectations, the market clearly did not expect recent price increases to be sustained in the long run. By overstating changes in expected gasoline prices, the martingale specification biases the model to expect larger changes in relative prices than the market should actually generate. This would bias the estimator in favor of concluding myopia; in the results section, we show that this bias can be severe.

6.4 Accounting for Elasticity on the Intensive Margin

Our primary specification of $G_{jat}$ described until now assumes that the elasticity of vehicle miles traveled with respect to gasoline prices is negligible. This section describes an alternative specification that accounts for two effects of this elasticity. First, changes in VMT change total expected gasoline expenditures $G_{jat}$. Second, the utility from vehicle ownership and use $\psi_{jat}$ also depends
on VMT: the utility from owning a vehicle and driving it 12,000 miles per year is different than
the utility of owning a vehicle and driving it 11,500 miles per year.

We adjust VMT for changes in gasoline prices using estimates of that elasticity from a recent
paper by Hughes, Knittel, and Sperling (2007). This analysis finds that between 2001 and 2006, the
short run elasticity of demand for vehicle miles traveled was between -.034 and -.077. We assume
an elasticity of -0.05 and translate this to a linear demand curve for computational ease, giving a
slope of $\frac{\partial m}{\partial g} = -444$ miles per dollar per gallon. We assume that the slope of this demand curve is
constant across all vehicles.

Figure 10.6 shows a vehicle owner’s demand for VMT. In 2001, gas prices are $g_{2001}$ and con-
sumers choose VMT $m_{ja,2001}(g_{2001})$, giving total annual gasoline costs in the shaded rectangle
bounded by those two values. At time $s$ with higher gasoline prices $g_s$, consumers reduce VMT to
$m_{jas}$. The new annual gasoline cost is now the unshaded rectangle bounded by $g_s/f_{jas}$ and $m_{jas}$. This approach is used to generate expected VMT at any possible gasoline price, using the following
formula:

$$m_{jas} = m_{ja,2001} + \frac{\partial m}{\partial g} \cdot (g_s - g_{2001}) \quad (19)$$

This adjusted VMT $m_{jas}$ is in turn used to generate an alternative value of $G_{jat}$.

The VMT demand curve also provides insight into how changes in VMT change consumers’
utility from vehicle use. The consumer’s total utility from vehicle use is the area under the demand
curve. As gasoline prices increase from $g_{2001}$ to $g_s$, this total utility decreases by the area of the
shaded trapezoid. Summing this over the future years of the vehicle’s life, we have an adjustment
denoted $I_{jat}$:

$$I_{jat} = \sum_{s=t+1}^{t+(L-1-a)} -f_{jas}^{-1} \cdot (g_{2001} + g_s) \cdot (m_{ja,2001} - m_{jas}) \cdot \phi_{jas} \cdot \beta^{s-t} \quad (20)$$

The sign of the variable $I_{jat}$ is defined such that as the utility from vehicle use decreases and
utility decreases, $I_{jat}$ increases. Because $I_{jat}$ is in measured in dollar terms, a one dollar increase in
$I_{jat}$ should reduce willingness to pay for the vehicle by one dollar. Because the slope of the VMT demand curve is extremely steep, it turns out that in practice $I_{jat}$ is highly collinear with $G_{jat}$. In our alternative specification that accounts for elasticity on the intensive margin, this means that we cannot include both $I_{jat}$ and $G_{jat}$ as independent variables. Instead, we move $I_{jat}$ to the left hand side and estimate equation (15) with $P_{jat} + I_{jat}$ as the dependent variable.

7 Results

We have estimated the reduced form equation (7) and nested logit equation (15) under a variety of assumptions. Throughout this section, we will refer to a base specification that includes observations from January 2003 to March 2008 of all passenger cars and light trucks age one to 25 years. In this specification, gasoline costs are computed according to equation (18) with a 15% discount rate, and expectations of future gasoline prices are determined using oil futures prices. This specification is relatively conservative in that we expect it to yield a higher estimate of consumer sensitivity to gasoline costs compared to other specifications. We choose this in order to avoid a false rejection of the hypothesis that vehicle prices fully adjust. The base specification uses the nested logit model with vehicle class as its only nest.

Table 3 shows the first stage regression results of the base model. The first stage can be viewed as a reduced form relation between new vehicle quantities and expected gasoline costs. Each column represents an instrumented variable. Logs of market share ($s_{jat}$) and within-group share ($s_{(ja/k)t}$) are regressed on expected gasoline costs ($G_{jat}$) and four instruments, all measured in $1,000s. The first two instruments are the expected gasoline costs in the year when the vehicle was new and the year before. The second two instruments are the respective means of the first two over all vehicles in the same nest. The negative coefficients on the first two instruments in column (1) suggest that auto makers produce fewer low fuel economy vehicles when gas price expectations are higher. The residuals in this regression are also shown graphically in Figure 10.7. The coefficients in column (2) are more difficult to interpret: each pair of instruments are correlated due to serial correlation in gas prices and futures prices, which likely explains the opposite signs of the coefficients. The first stage estimated without expected gas costs before the vehicle was produced yields the expected
sign on the coefficients: higher expected gas costs are associated with lower within-nest market shares, while higher mean expected gas costs across a nest are associated with larger within-nest market shares. This result is qualitatively consistent with analyses of new vehicle supply in recent analyses by Klier and Linn (2008), Li, Timmins, and von Haefen (2009), and Busse, Knittel, and Zettelmeyer (2009).

Table 4 compares our approaches to estimating \( \hat{\gamma}/\eta \). Column (1) shows an estimation of the reduced form equation (7) in which quantities are assumed to be constant. The coefficient on gas cost is the negative of the estimate of \( \gamma/\eta \), so that \( \hat{\gamma}/\eta = 0.08 \). Under the reduced form assumptions, this suggests that automobile consumers account for only 8% of gasoline costs when purchasing a vehicle. The residuals are plotted in Figure 10.8. Column (2) shows an estimation of a logit specification with no substitution patterns. The point estimate of \( \hat{\gamma}/\eta \) is triple that of the reduced form, which is consistent with our framework that ignoring the response of market shares to gasoline prices will attenuate the apparent response of vehicles prices to expected gasoline cost changes. The coefficient on market share of -4,120 suggests that a one percent increase in the market share of a vehicle causes a $41 drop in its price. Column (3) shows the base nested logit model, estimated using ordinary least squares. The market share coefficient drops sharply, likely due to correlation between market shares and unobserved vehicle quality. \( \hat{\gamma}/\eta \) is slightly negative here, making the result difficult to interpret. Column (4) shows an instrumental variables estimation of the base nested logit model, and Figure 10.9 displays the second stage graphically. \( \hat{\gamma}/\eta \) and \( \hat{\gamma}/\eta \) are similar to those in the logit model. The coefficients on log market share and log nest share suggest that the correlation parameter \( \sigma \) is about 0.1, so that utilities within a class are not highly correlated. This is consistent with the small difference between the logit and nested logit specifications. Our base specification suggests that consumers account for 25% of gasoline costs.

If the nest structure does not adequately capture substitution patterns, we may overstate the willingness of consumers to switch vehicles as gas price expectations change, biasing \( \hat{\gamma}/\eta \) towards zero. Table 5 explores alternative nest structures, starting with a two level nest structure in column (2). Utilities are allowed to be correlated within a vehicle class, but an additional correlation is allowed among vehicles in the same vehicle class and age category. We define the age categories to be 0-5 years, 6-10 years, and greater than 10 years, since Stolyarov (2002) suggests that consumers
are most likely to trade in vehicles around age 5 or 10 in order to purchase a newer one. In this specification, a correlation is not allowed among vehicles of the same age category but of different vehicle classes. Column (3) switches the order of these nests, in case this substitution within age groups is relatively more important. Column (4) uses a two level nested logit where the first nest is vehicle class and the second nest is the interaction of an indicator for whether the vehicle is of a luxury make and an indicator for the continent where the firm is based (Europe, North America, or Asia). This captures preferences of consumers to purchase a certain “style” of vehicle, such as a European luxury mid-size sedan. Column (5) includes three nests: class first, age category second, and style third. For models with two nests, ‘nest 1 share’ denotes the share of all vehicles in the same second nest within the first nest, and ‘nest 2 share’ denotes the share of a vehicle within the second nest. Nest shares are similarly defined for the model with three nests. The respective coefficients on each logged share are the $\sigma$ parameters related to the correlations of utilities within that nest. While the coefficients on share variables change substantially as the nest structure is changed, response to gas prices is changed little. While these alternative nest structures does not exhaust all possible forms of substitution patterns, they do suggest that uncaptured substitution patterns in the base model are unlikely to cause a bias in $\frac{\gamma}{\eta}$.

Our model makes an implicit assumption about the functional form by which consumers compute a discounted sum of expected future gasoline costs. Alternate assumptions are that gasoline costs are not weighted, ($\frac{\gamma}{\eta} = 1$) but instead consumers use a discount rate that may not correspond to market interest rates, or that consumers have a limited time horizon and only account for a fixed number of years of gasoline costs. In principle, we could differentiate between these, but in practice, we do not have enough power in our estimation to identify the entire functional form of the discounting. However, it is instructive to test whether $\frac{\gamma}{\eta} = 1$ under different assumptions about consumers discounting the future. Columns (2)-(4) in Table 6 show estimates of $\frac{\gamma}{\eta}$ when consumers use a 5%, 25%, and 45% annual discount rate, respectively, rather than 15% in the base specification (column 1). As one would expect, consumers appear more sensitive to gasoline costs when the discount rate is higher. The confidence interval on $\frac{\gamma}{\eta}$ includes one only when the discount rate reaches an implausibly high 45%. Column (5) repeats the base specification, but gas costs are only added for the first three years. Even with a short time horizon, consumers appear to account
for only 43% of gas costs. Therefore, the observed undersensitivity to gas costs cannot be explained by consumers who plan to resell their vehicle after a relatively short period but do not expect the resale price to depend on gas price expectations.

Consumers’ expectations of future gasoline prices and vehicle usage are central to the definition of $G_{jat}$, and in particular, overstating expected future gas costs would lead to a downward bias in $\frac{\hat{\sigma}}{\eta}$. In column (2) of Table 7, future gas costs are calculated only for years in which futures data are available to make a prediction. Effectively, consumers are assumed to ignore gas costs after this point. Estimates here are comparable to the base specification. In column (3), gas prices are assumed to be martingale, so that any change in spot gas prices is assumed to be permanent. If consumers indeed believed that gas prices are martingale, they should be very responsive to current prices. It is therefore not surprising that $\frac{\hat{\sigma}}{\eta}$ is much smaller than for the base specification, where short-term changes are assumed to revert to the mean over time. In column (4), $G_{jat}$ is computed assuming that all vehicles have a lifetime of 13 years (the expected lifetime of a new vehicle in our data set) and are driven 12,000 miles each year. We again assume martingale prices. The fixed lifetime assumption requires us to drop vehicles with age greater than 13 years. However, since the fixed lifetime places strong assumptions on vehicles near age 13, we limit the dataset to vehicles age 8 or less. In column (5), we allow consumers to adjust their Vehicle Miles Traveled (VMT) as gas prices change. We also add an adjustment to the price as compensation for the reduced use of the vehicle, as described in section 6.4. Neither of these assumptions alters $\frac{\hat{\sigma}}{\eta}$ substantially.

In column (2) of Table 8, we use the same specification but include data through the end of 2008. The large changes in the automobile market and the economy in general after March 2008 may have changed consumer preferences substantially, and these preference changes are not included in our model. However, the large change in gas prices and gas price expectations may help to identify the parameter of interest. Including this time period makes our estimate of $\frac{\hat{\sigma}}{\eta}$ substantially larger, just over one half. While still short of full adjustment, these price changes may suggest that consumers were much more responsive to gas prices during this time period, if any changes in preferences did not differentially affect demand for high and low MPG vehicles. One explanation for this finding is that consumers are undersensitive to future gas costs when gas prices are below some “threshold,” perhaps between $3 and $4 per gallon. The gas prices spike in mid 2008 may have led to a sudden
adjustment in vehicle prices that is not observed with smaller changes in gas price expectations.

7.1 Magnitude of the Fuel Economy Discount

What does incorporating 25% of gasoline prices mean for vehicle prices? Does this mean that used vehicles are substantially mis-priced, or are these in practice only small deviations from the equilibria that we would expect to see?

Our empirical approach is not informative about the absolute mispricing of different vehicles. Our coefficient estimates do predict, however, how much the relative prices and quantities of vehicles with different fuel economy ratings will change in response to a given change in gasoline price expectations. The "Predicted Price Change" line in Figure 10.10 presents these changes for example used vehicles that will be driven 12,000 miles per year for the remaining seven years of their lifetimes, given a permanent $1 increase in gasoline prices. This line is computed by first fitting the expected change in vehicle market shares using the first stage coefficients. We then take those fitted market share changes, along with the change in $G_{jat}$, and use our second stage coefficients to predict the change in vehicle prices given the $\frac{\hat{e}}{\eta}$ from our primary specification. The Honda Civic, which has fuel economy of 37 miles per gallon and thus 0.027 gallons per mile, is normalized to have zero price change.

The graph shows that our model predicts that a Ford F-150 with the same expected VMT and lifetime, which requires 0.050 gallons of gasoline per mile, would see its relative price drop by $180. The double line on the graph presents the relative price changes that would be expected if $\frac{\hat{e}}{\eta} = 1$. For the F-150, again holding constant the changes in market shares predicted by the first stage, the relative price should drop by $1050 instead of $180.

A simple back of the envelope calculation can be used to understand the order of magnitude of the "Fuel Economy Discount Puzzle" across the entire US vehicle market. There are approximately 200 million passenger vehicles registered in the United States, and this example F-150 with seven years of remaining life is very close to the mean vehicle in the US vehicle fleet in terms of gallons per mile and remaining life. Our model predicts that the relative price of this mean vehicle under-adjusted by nearly $800. The relative price of this fleet of vehicles, therefore, should have adjusted
by $160 billion more than it did in response to a permanent $1/gallon increase in expected gasoline prices. This is on the order of 5 percent of the total value of the US passenger vehicle fleet.

8 Conclusion

This paper investigates whether vehicle prices and market shares respond to changes in gasoline price expectations in a way that is consistent with rational, forward-looking consumers who value $1 today the same as they value future spending with present discount value of $1. We use a discrete choice framework to model consumer demand and substitution patterns and show that neglecting these effects would attenuate the estimated vehicle price response of interest. The price and quantity changes strongly suggest that consumers substantially underweight future gas costs – depending on the specification, consumers are willing to pay only $0.25 up front to reduce future gasoline expenditures by $1 (in present discounted value terms). The estimated responsiveness of vehicle prices to expected gas costs depends on assumptions about expectations of future gas prices, vehicle lifetime, consumer discount rates, substitution patterns, and other parameters. However, we find that prices fully adjust only under implausible sets of assumptions.

The finding that consumers underweight gas costs suggests that a gasoline tax increase or a carbon tax intended to offset negative externalities would have less than the efficient level of effects on consumers’ vehicle choice. This may provide economic justification for the use of fuel economy standards as part of climate change policy in the United States. Our estimation of vehicle demand, along with a model of new vehicle supply, could be used to predict the effect of changes in gasoline price expectations on vehicle prices and quantities of new vehicles. This framework would allow a comparison of the impacts of a gasoline tax increase or a fuel economy standard on the distribution of new vehicle fuel economy, gasoline consumption, and consumer welfare.

Our finding also raises the questions of why consumers underweight future gasoline costs, and whether the magnitude of apparent myopic behavior is correlated with observable characteristics. Future work that explains which consumers are less sensitive to future gasoline costs more may suggest a more targeted policy to correct this behavior.
References


9 Appendices

9.1 Deriving the Utility Function

In this Appendix, we derive our static discrete choice model from a more realistic model of the consumer's decision problem. In the process, we clarify and discuss the assumptions required for our estimator to be consistent.

We build on the approach of Stolyarov (2002) in writing down the consumer’s dynamic durable goods choice problem. The consumer maximizes an indirect utility function \( U = \eta w + u_{ijat}(\eta, \gamma, \lambda) \), which is additively separable in "vehicle utility" \( u_{ijat} \) and consumption of a numeraire good. As in the text, consumer \( i \) chooses a vehicle in period \( t \) from the set \( J,A \) of possible model-by-age combinations. Owning vehicle \( ja \) at time \( t \) forces expected one-period gasoline expenditures \( \bar{G}_{jat} \) and gives one-period utility flow \( \bar{\psi}_{ijat} \). This individual-specific utility flow is the sum of average utility \( \bar{\psi}_{jat} \) and an individual taste error \( \bar{e}_{ijat} \). In the next year, where utility flows are discounted by factor \( \beta \), the consumer will have the choice to sell the vehicle, incurring transaction cost \( \lambda_{ja} \), or hold it.

In the body of the paper, we assumed risk neutrality, homogeneous \( \gamma \) and \( \eta \), and that \( G \) does not vary within-\( jat \). Under these assumptions, the consumer maximizes the following Bellman Equation:

\[
\max_{J,A} u_{ijat} = -\eta p_{jat} - \gamma \bar{G}_{jat} + \bar{\psi}_{ijat} + \beta \max \left\{ \max_{A,K} \left\{ u_{ikat+1} + \eta p_{j(a+1)(t+1)} + \lambda_{ja} \right\} + \eta p_{j,a+1,t+1} + \eta p_{j,a+1,t+1} \right\} \quad (21)
\]

Most analyses of durable goods markets, including Stolyarov (2002) and the literature following Rust (1985), assume that the market is stationary: the prices, quantities, and attributes of the choice set remain constant. This is useful for us because it prevents us from needing to make a series of other, potentially more complex assumptions about how consumers believe the market will evolve.
**Assumption: Stationarity:** \( E[p_{jat+s}] = p_{jat} \) and \( E[s_{jat+s}] = s_{jat}, \forall s \) and \( E[J_{A_{t+s}}] = J_{A_t} \)

Stolyarov (2002) shows that if the market is stationary, the consumer’s decision rule is also stationary: she will purchase her preferred vehicle, hold that preferred vehicle as it ages until the utility gain from replacing the vehicle outweighs the transaction cost \( \lambda_{ja} \), and then replace with the same preferred vehicle. We denote the optimal holding period as \( \tau_{jat} \). For expositional ease, this is assumed to be constant within the set of consumers that purchase vehicle \( ja \) at time \( t \) in equilibrium. The vehicle utility from buying vehicle \( ja \) is:

\[
0_{ijat} = -\eta p_{jat} + \sum_{s=0}^{\tau-1} \beta^s \left( \gamma G_{j,a+s,t+s} + \tilde{\psi}_{i,j,a+s,t+s} \right) \\
+ \beta^\tau \left( \eta p_{j,a+t+\tau} - \lambda_{ja} \right) + \beta^\tau 0_{ija(t+\tau)} \tag{22}
\]

The first line captures the utility from paying for the vehicle and then fueling and using it over \( \tau \) years. The first term on the second line captures the discounted utility from selling it, including the transaction cost. The last term reflects the fact that, in a stationary market, the consumer will re-purchase the same vehicle - and realize the same utility - over the next \( \tau \) years.

We can be more specific about the resale price by assuming that consumers expect the prices predicted by the nested logit model. Recall that this gives:

\[
p_{jat} = \frac{1}{\eta} \left[ -(\ln s_{jat} - \ln s_{0t}) - \gamma G_{jat} + \sigma_t \ln s_{(j/l)at} + \psi_{jat} \right] \tag{23}
\]

Substituting this into the utility function, we have:

\[
0_{ijat} = -\eta p_{jat} + \sum_{s=0}^{\tau-1} \beta^s \left( \gamma G_{j,a+s,t+s} + \tilde{\psi}_{i,j,a+s,t+s} \right) \\
+ \beta^\tau \left( -(\ln s_{j,a+t+\tau} - \ln s_{0,t+\tau}) + \sigma_t \ln s_{(j/l)at} + \psi_{j,a+\tau+t+s} \right) + \sum_{s=0}^{L-1-a} \beta^s \left( \gamma G_{j,a+s,t+s} + \tilde{\psi}_{j,a+s,t+s} \right) - \lambda_{ja} \right) \\
+ \beta^\tau 0_{ija(t+\tau)} \\
= -\eta p_{jat} + \sum_{s=0}^{L-1-a} \beta^s \left( \gamma G_{j,a+s,t+s} + \tilde{\psi}_{i,j,a+s,t+s} \right) + \sum_{s=0}^{\tau-1} \tilde{e}_{ij,a+s,t+s} \\
+ \beta^\tau \left( -(\ln s_{j,a+t+\tau} - \ln s_{0,t+\tau}) + \sigma_t \ln s_{(j/l)at} + \psi_{j,a+\tau+t+s} - \lambda_{ja} \right) + \beta^\tau 0_{ija(t+\tau)} \tag{24}
\]

The first term of the last line reflects that part of the vehicle’s resale value depends on future market share. The last term generates an infinite sum of utilities. We now omit it, as it simply scales nominal utility by an amount that depends on the consumer’s time horizon. Note that the introduction of a constant vehicle death probability is simply equivalent to decreasing the discount factor.

By specifically defining some of the terms from our apparently-static utility function, we can now show that our dynamic model maps into this utility function. This allows us to make explicit the assumptions required for our estimator to be consistent in a dynamic world. Recall that our apparently-static utility function was:
\[ u_{ijat} = -\eta p_{jat} - \gamma G_{jat} + \psi_{jat} + \varepsilon_{ijat} \]

We map the "dynamic" variables into the "static" variables with the following equations:

\[
G_{jat} = \sum_{s=0}^{L-1} \beta^s \tilde{G}_{j,a+s,t+s}
\]

\[
\psi_{jat} = \sum_{s=0}^{L-1} \beta^s \tilde{\psi}_{j,a+s,t+s} + \beta^\tau \left( -\ln s_{j,a+\tau,t+\tau} - \ln s_{0,t+\tau} + \sigma \ln s_{(j,t)at} - \lambda_{ja} \right)
\]

\[
\varepsilon_{ijat} = \sum_{s=0}^{T-1} \beta^s \tilde{\varepsilon}_{ij,a+s,t+s}
\]

The first line indicates that, as before, we can define \(G_{jat}\) as the discounted sum of future fuel costs. The second line now captures both the consumer’s utility from using the vehicle and the resale value and transaction cost. The third line is the individual-specific error term, which we assume takes the "nested logit" structure.

### 9.2 Data

In this data appendix, we describe in detail the construction and cleaning of vehicle price and quantity data, vehicle attributes, and future expected gasoline prices. We then detail how the data from multiple sources was merged into one dataset.

#### 9.2.1 Vehicle Price and Quantity Data

**Vehicle Prices**  The Manheim dataset consists of observations of individual vehicles put up for sale at a Manheim auction. We keep observations that resulted in a sale and for which we have a valid VIN number that matches our other data sets. Prices are adjusted for inflation, logged, then used as the left hand side variable in a fixed effects regression containing odometer reading and its square, dummies for vehicle condition code, region of sale, type of sale (open to the public or restricted to certain buyers), and auction type (physical in-lane auction or internet sale). The fixed effects are model by model year by year by month. One logged price is predicted for each fixed effect, assuming a vehicle with an odometer reading predicted using the NHTS data, in ‘good’ condition, sold in the Midwest, in an open sale, in a physical auction. These predicted values are then exponentiated to obtain monthly price estimates for that model and model year.

**Quantities**  Vehicle quantities are annual snapshots of registration data collected for all new and used vehicles in the entire United States by R.L. Polk. We assume that the quantity in any month is equal to the registered quantity in the July snapshot. Since registrations are typically renewed annually or biennially, there may be slight differences between the registration snapshots and the actual quantities of a model in use at the time. New vehicles are a particular problem in that not all vehicles are registered by July of the model year. Since very few vehicles are retired in the first few years after the model year, we set the quantity in the model year equal to the quantity one year later.

---

22 The data preparation for this project was unusually complex and time consuming. The bulk of it was undertaken by a very detail-oriented research assistant named Sounman Hong. We thank him again for his efforts.
9.2.2 Vehicle Attributes

**Fuel Economy**  Since 1975, the EPA has employed a consistent test, called a dynamometer test, to measure fuel economy. In 1985, the EPA introduced adjustment factors to these tests to account for an "in-use shortfall," the difference between fuel economy computed under laboratory conditions and the actual fuel economy that the EPA measured for drivers on the road. The "Adjusted" values were computed for City and Highway MPG by multiplying the Laboratory values by 0.9 and 0.78, respectively, and these Adjusted values were the ones made public for consumers for model years 1985-2007.

During the past several years, the EPA has again adjusted its fuel economy calculation to account for changes in driving patterns since 1984. For model year 2005-2008 vehicles, these New Adjusted values are:

- New Adjusted City = $1/(0.003259+1.1805/Lab City)$
- New Adjusted Highway = $1/(0.001376+1.3466/Lab Highway)$

The EPA believes that its previous adjustment factors were valid for model years 1975-1985. For model years 1986-2005, however, the EPA now retroactively assumes that these New Adjusted fuel economy values phased in linearly over the 20 year period.

To construct a Composite fuel economy rating from the reported MPG’s between 1975 and 2007, inclusive, the EPA originally took the weighted harmonic mean of City and Highway New Adjusted MPG ratings, with 55% and 45% weights, respectively. In its 2008 revision, however, the EPA changed these weights to 43% and 57%, respectively. The EPA now retroactively assumes that this change actually occurred linearly between 1986 and 2005.

We construct two measures of fuel economy, one which should reflect consumers’ best guess at MPG based on information publicly available at the time, and one which reflects analysts’ best guess in 2008 at what each vehicle’s fuel economy actually was. We use the former in our primary specification, to reflect what consumers would have believed.

For our 2008 adjusted MPG, we use the 0.9 and 0.78 Adjustment factors before 1985. For 1985-2008, we use the (retroactively phased in) New Adjusted EPA methodology. Finally, Greene, et al, (2007) report that fuel economy in used cars degrades at an average of 0.07 MPG per year. We further adjust MPG to account for this.

**Other Attributes**  We use data on vehicle characteristics, including horsepower, weight, and wheelbase. For all model years, these data are from the Ward’s Automotive Yearbook. These were purchased in electronic form from Ward’s for model years beginning with 1995. We use curb weight as the measure of weight.

**Vehicle Class**  Vehicle class data is from the EPA when available. When EPA data is not available, we use vehicle characteristics to determine vehicle class consistent with EPA’s definition. Cars are divided into two-seaters (which seat only two adults) and sedans, which are further subdivided into minicompact, subcompact, compact, mid-size, and large based on interior volume. Trucks are divided into pickup trucks, sport utility vehicles, minivans, and vans based on their intended purpose. Pickup trucks and SUVs are further subdivided into standard and small based on gross vehicle weight rating, but we ignore this distinction, as vehicles may be highly substitutable across these categories. An additional class of light trucks, special purpose vehicle, is not used in recent years but includes pickup trucks, SUVs, and minivans. These are manually recoded into the most appropriate class.

9.2.3 Future Gasoline Costs

Computation of discounted future gasoline costs $G_{jat}$ requires the expected gas price, expected vehicle miles traveled, and probability that the vehicle is still functional for all future time periods. We outline the computation of each of these.

---

23 The information in this and the following paragraph is from the U.S. EPA (2008). We received the EPA adjusted and unadjusted test data for 1975-2008 directly from researchers at Oak Ridge National Laboratory.
Inflation Adjustment  All vehicles prices and gasoline prices are deflated using the BLS consumer price index series for All Items, Urban (CUUR0000SA0), available from ftp://ftp.bls.gov/pub/time.series/cu/cu.data.1.AllItems. We use real dollars for the average CPI in 2005, which is equivalent to real July 2005 dollars.

Gasoline Price  Our source of gasoline price data is the US Energy Information Administration (EIA). We use US city average retail gasoline price for all types of gasoline, which are available on a monthly basis from Table 9.4 of the EIA’s Monthly Energy Review.

A vehicle of a given model year typically begins being sold in September of the previous calendar year. The monthly gasoline prices used to construct the instrument were thus the September-August mean gasoline prices.

Survival Probability  In our base specification, the probability that a vehicle is still functional in a future time period is estimated using a probit model with grouped data. The outcome variable is the number of vehicles of a model and model year registered at time \( t \), \( q_{jat} \), out of the number of vehicles registered when that vehicle was new, \( q_{j0(t-a)} \), from the R.L. Polk data. The estimation of survival probabilities recovers coefficients on age dummies, model year and its square, vehicle class dummies, firm dummies, and firm-specific linear age trends. The sample used in the estimation is the same as the sample used in our discrete choice model, except that vehicles for which quantity is unknown at age zero are dropped. The estimation coefficients are used to predict the probability that each vehicle in the data set survives to at least \( s \) years beyond its current age, for all positive \( s \) up to \( L = 25 \). The probability that each vehicle survives to its current age is also predicted. The former is divided by the latter to obtain the probability that a vehicle survives \( s \) more years conditional on surviving to its current age. This is the relevant probability that enters the computation of \( G_{jat} \). Due to scarce data for vehicles older than 25 years, we set the probability that a vehicle survives past age 25 equal to zero. The trends in the data set suggest this is not an unreasonable assumption.

Vehicle Miles Traveled  We do not observe average Vehicle Miles Traveled (VMT) for all vehicles on the road. Instead, we use microdata from the National Household Travel Survey (NHTS) for 2001. These data allow us to predict the expectation of a vehicle’s VMT conditional on its characteristics. There are two possible measures of VMT included in the data: consumers’ stated VMT and recorded odometer readings. Because we are interested in consumers’ expectations of VMT, we use the stated VMT.

9.2.4 Data Construction and Coverage

Our data are merged by partial VIN number using the Complete Prefix File, a product sold by R.L. Polk. This allows us to use a common set of vehicle names and descriptions throughout the data set. Wards and EPA data do not contain VIN information, so these were matched by name. Each dataset provides information at different levels of detail: one dataset may include separate information for a two wheel drive versus a four wheel drive version of a model, while another includes only mean information on that model. We have collapsed the dataset to the most disaggregated level that is feasible given the data constraints. In this collapsing process, prices are estimated with number of observations as weights, quantities are summed, MPG is averaged using the harmonic mean, and other characteristics are averaged using an arithmetic mean.

The aim of the dataset is to include consumers’ entire vehicle choice sets for every month between 2003 and 2008. This includes all light duty vehicles (cars and light trucks) available to the public. Due to data constraints, we had to drop parts of the choice set; this is not uncommon in discrete choice models where data on small parts of the choice set may not be available. In particular, we dropped vehicles for which we are missing one or more of the required data sources: prices, quantities, or fuel economy. Vehicles with model years before 1983 were also dropped, as we can only match VIN numbers to a vehicle name after 1983. Furthermore, we drop vehicles which do not use gasoline or are clearly not part of a typical consumer’s choice set, such as delivery vehicles and motor homes.
10 Figures

10.1 Vehicle Prices

10.2 Vehicle Quantities
10.3 Fuel Economy Ratings of Vehicles Registered in 2007

10.4 Vehicle Miles Traveled By Vehicle Age
10.5 Gasoline Price Expectations

![Graph showing Gasoline Spot and Oil Futures Prices]

10.6 Intensive Margin

![Diagram illustrating Intensive Margin]

\[ P \]  
\( \text{($) / mile) } \)

\[ \frac{g_j}{f_{jas}} \]

\[ I_{jat} \]

\[ \frac{g_{2001}}{f_{jas}} \]

\[ m_{jas} \]

\[ m_{ja,2001} \]  

VMT (miles)
10.7 First Stage

Because of a problem in Stata, the coefficient and standard error estimates do not correspond to our coefficient estimates in the tables above.

10.8 Reduced Form

Because of a problem in Stata, the coefficient and standard error estimates do not correspond to our coefficient estimates in the tables above.
10.9 Primary Specification

Because of a problem in Stata, the coefficient and standard error estimates do not correspond to our coefficient estimates in the tables above.

10.10 Predicted Mispricing with a $1/Gallon Gas Price Increase

Note: Price levels have been normalized such that a vehicle with 0.027 gallons per mile (the Honda Civic) has zero mispricing. The figure indicates predicted price changes relative to this vehicle if gasoline prices increase by $1/gallon.
# 11 Tables

<table>
<thead>
<tr>
<th>Table 1: <strong>Summary statistics</strong></th>
<th>Full sample</th>
<th>2005 new models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
<td>2005.6</td>
<td>2005.0</td>
</tr>
<tr>
<td></td>
<td>(1.7)</td>
<td>(0.0)</td>
</tr>
<tr>
<td><strong>Model Year</strong></td>
<td>1997.6</td>
<td>2005.0</td>
</tr>
<tr>
<td></td>
<td>(5.9)</td>
<td>(0.0)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td>8.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>(5.7)</td>
<td>(0.0)</td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td>7,685</td>
<td>24,778</td>
</tr>
<tr>
<td></td>
<td>(8,757)</td>
<td>(9,424)</td>
</tr>
<tr>
<td><strong>Miles per gallon</strong></td>
<td>19.6</td>
<td>19.9</td>
</tr>
<tr>
<td></td>
<td>(4.2)</td>
<td>(4.8)</td>
</tr>
<tr>
<td><strong>Expected lifetime gas costs</strong></td>
<td>3,312</td>
<td>4,663</td>
</tr>
<tr>
<td>(2005 $)</td>
<td>(2,142)</td>
<td>(1,805)</td>
</tr>
<tr>
<td><strong>Horsepower</strong></td>
<td>199.2</td>
<td>210.7</td>
</tr>
<tr>
<td></td>
<td>(57.0)</td>
<td>(60.3)</td>
</tr>
<tr>
<td><strong>Weight (pounds)</strong></td>
<td>3,741</td>
<td>3,998</td>
</tr>
<tr>
<td></td>
<td>(1,025)</td>
<td>(1,913)</td>
</tr>
<tr>
<td><strong>Wheelbase (inches)</strong></td>
<td>113.4</td>
<td>114.3</td>
</tr>
<tr>
<td></td>
<td>(14.4)</td>
<td>(14.0)</td>
</tr>
<tr>
<td><strong>Fraction cars</strong></td>
<td>0.58</td>
<td>0.48</td>
</tr>
</tbody>
</table>

**Observations** 1,221,795 9,934

Notes: Quantity-weighted means are shown with standard deviations in parenthesis below. The full sample includes monthly observations from January 2003 to December 2008 of all passenger cars and light trucks age 0-25. Column (2) includes 2005 model year vehicles observed in 2005. See text for calculation of expected gas costs.
Table 2: Gasoline Prices and Expectations

<table>
<thead>
<tr>
<th>Year</th>
<th>Spot</th>
<th>Future Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-1</td>
<td>1-2</td>
</tr>
<tr>
<td>1998</td>
<td>1.34</td>
<td>1.44</td>
</tr>
<tr>
<td>1999</td>
<td>1.43</td>
<td>1.50</td>
</tr>
<tr>
<td>2000</td>
<td>1.77</td>
<td>1.73</td>
</tr>
<tr>
<td>2001</td>
<td>1.69</td>
<td>1.65</td>
</tr>
<tr>
<td>2002</td>
<td>1.56</td>
<td>1.63</td>
</tr>
<tr>
<td>2003</td>
<td>1.74</td>
<td>1.71</td>
</tr>
<tr>
<td>2004</td>
<td>1.99</td>
<td>1.95</td>
</tr>
<tr>
<td>2005</td>
<td>2.34</td>
<td>2.33</td>
</tr>
<tr>
<td>2006</td>
<td>2.55</td>
<td>2.55</td>
</tr>
<tr>
<td>2007</td>
<td>2.68</td>
<td>2.59</td>
</tr>
<tr>
<td>2008</td>
<td>3.00</td>
<td>3.12</td>
</tr>
</tbody>
</table>

Table 3: Nested logit first stage

<table>
<thead>
<tr>
<th>Instrumented variable:</th>
<th>(1) ln ( s_{jat} )</th>
<th>(2) ln ( s_{(ja/k)t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G_{jat} )</td>
<td>-0.019 ( 0.020 )</td>
<td>0.014 ( 0.002 )</td>
</tr>
<tr>
<td>( G ) at age 0</td>
<td>-0.030 ( 0.027 )</td>
<td>-0.004 ( 0.002 )</td>
</tr>
<tr>
<td>( G ) at age -1</td>
<td>-0.064 ( 0.028 )</td>
<td>0.001 ( 0.002 )</td>
</tr>
<tr>
<td>Mean(( G ) at age 0)</td>
<td>0.297 ( 0.158 )</td>
<td>1.464 ( 0.014 )</td>
</tr>
<tr>
<td>Mean(( G ) at age -1)</td>
<td>-0.460 ( 0.225 )</td>
<td>-1.623 ( 0.032 )</td>
</tr>
<tr>
<td>Observations</td>
<td>546,014</td>
<td>546,014</td>
</tr>
<tr>
<td>ja groups</td>
<td>24,665</td>
<td>24,665</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.95</td>
<td>0.91</td>
</tr>
<tr>
<td>F (excl instruments)</td>
<td>4.65</td>
<td>4575.24</td>
</tr>
</tbody>
</table>

Notes: The sample includes monthly observations from January 2003 to March 2008 of all passenger cars and light trucks age 1-25. See text for calculation of expected gas costs. The excluded instruments are the last four variables and are described in detail in the text. ‘\( G \) at age 0’ and ‘\( G \) at age -1’ are the expected gas costs in the year the vehicle was new and one year earlier. The second two instruments are the mean of these quantities within a vehicle’s class. All regressors are measured in $1,000s. Standard errors are robust and clustered by \( ja \) (model * age).
Table 4: Nested logit first stage

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrumented variable:</td>
<td>ln $s_{jat}$</td>
<td>ln $s_{(ja/k)t}$</td>
</tr>
<tr>
<td>$G_{jat}$</td>
<td>-0.021</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>G at age 0</td>
<td>-0.065</td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Mean(G at age 0)</td>
<td>0.155</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Observations</td>
<td>546,014</td>
<td>546,014</td>
</tr>
<tr>
<td>ja groups</td>
<td>24,665</td>
<td>24,665</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.05</td>
<td>0.74</td>
</tr>
<tr>
<td>F (excl instruments)</td>
<td>3.98</td>
<td>2045.64</td>
</tr>
</tbody>
</table>

Notes: The sample includes monthly observations from January 2003 to March 2008 of all passenger cars and light trucks age 1-25. See text for calculation of expected gas costs. The excluded instruments are the last four variables and are described in detail in the text. ‘G at age 0’ and ‘G at age -1’ are the expected gas costs in the year the vehicle was new and one year earlier. The second two instruments are the mean of these quantities within a vehicle’s class. All regressors are measured in $1,000s. Standard errors are robust and clustered by ja (model * age).

Table 5: Comparison of reduced form and nested logit

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reduced form</td>
<td>Logit</td>
<td>Nested Logit</td>
<td>OLS</td>
</tr>
<tr>
<td>$G_{jat}$</td>
<td>-0.08</td>
<td>-0.28</td>
<td>0.01</td>
<td>-0.25</td>
</tr>
<tr>
<td>$[-\gamma/\eta]$</td>
<td>(0.04)</td>
<td>(0.10)</td>
<td>(0.04)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>ln(market share)</td>
<td>-4,120</td>
<td>-1,034</td>
<td>-4,561</td>
<td></td>
</tr>
<tr>
<td>$[-1/\eta]$</td>
<td>(1,548)</td>
<td>(329)</td>
<td>(1,588)</td>
<td></td>
</tr>
<tr>
<td>ln(nest share)</td>
<td>1,160</td>
<td>410</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$[\sigma/\eta]$</td>
<td>(328)</td>
<td>(581)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>546,014</td>
<td>546,014</td>
<td>546,014</td>
<td>546,014</td>
</tr>
<tr>
<td>ja groups</td>
<td>24,665</td>
<td>24,665</td>
<td>24,665</td>
<td>24,665</td>
</tr>
</tbody>
</table>

Notes: The sample includes monthly observations from January 2003 to March 2008 of all passenger cars and light trucks age 1-25. See text for calculation of expected gas costs. All specifications include monthly time dummies. Market share variables are instrumented in columns (2) and (4). Columns (3) and (4) use a nested logit model with vehicle class as the only nest. Nest share denotes the share of a vehicle within its class. Standard errors are robust and clustered by ja (model * age).
Table 6: Alternative nest structures

<table>
<thead>
<tr>
<th>Specification</th>
<th>(1) Base</th>
<th>(2) Class/Age</th>
<th>(3) Age/Class</th>
<th>(4) Class/Style</th>
<th>(5) 3 nests</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{jat}$</td>
<td>-0.25</td>
<td>-0.28</td>
<td>-0.25</td>
<td>-0.22</td>
<td>-0.24</td>
</tr>
<tr>
<td>$[-\gamma/\eta]$</td>
<td>(0.11)</td>
<td>(0.08)</td>
<td>(0.05)</td>
<td>(0.11)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>ln(share)</td>
<td>-4,561</td>
<td>-3,069</td>
<td>-8,790</td>
<td>-4,929</td>
<td>-2,149</td>
</tr>
<tr>
<td>$[-1/\eta]$</td>
<td>(1,588)</td>
<td>(1,206)</td>
<td>(1,509)</td>
<td>(1,964)</td>
<td>(909)</td>
</tr>
<tr>
<td>ln(nest 1 share)</td>
<td>410</td>
<td>-3,247</td>
<td>7,161</td>
<td>-6,652</td>
<td>-3,086</td>
</tr>
<tr>
<td>$[\sigma_1/\eta]$</td>
<td>(581)</td>
<td>(1,129)</td>
<td>(2,539)</td>
<td>(6,216)</td>
<td>(903)</td>
</tr>
<tr>
<td>ln(nest 2 share)</td>
<td>-7</td>
<td>7,552</td>
<td>1,348</td>
<td></td>
<td>-3,334</td>
</tr>
<tr>
<td>$[\sigma_2/\eta]$</td>
<td>(515)</td>
<td>(2,294)</td>
<td>(969)</td>
<td></td>
<td>(1,800)</td>
</tr>
<tr>
<td>ln(nest 3 share)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>516</td>
</tr>
<tr>
<td>$[\sigma_3/\eta]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(490)</td>
</tr>
<tr>
<td>Observations</td>
<td>546,014</td>
<td>546,014</td>
<td>546,014</td>
<td>546,014</td>
<td>546,014</td>
</tr>
<tr>
<td>ja groups</td>
<td>24,665</td>
<td>24,665</td>
<td>24,665</td>
<td>24,665</td>
<td>24,665</td>
</tr>
</tbody>
</table>

Notes: The sample includes monthly observations from January 2003 to March 2008 of all passenger cars and light trucks age 1-25. See text for calculation of expected gas costs. All specifications include monthly time dummies. Column (1) matches the last column of Table 3 and shows a nested logit model with vehicle class as the only nest. Here, nest 1 share denotes the share of a vehicle within its class. Column (2) uses two nests, vehicle class, and age buckets (0-4 years, 5-10 years, 11+ years). Nest 1 share denotes the share of a vehicle’s age and class nest within its class nest, and nest 2 share denotes the share of a model within its class and age nest. Column (3) reverses the order of the nests. Column (4) uses as a second nest the “style” of a vehicle: an indicator for whether the vehicle is a luxury make interacted with indicators for whether the firm is based in Europe, North America, or Asia. Column (5) includes all three nests, and nest 3 share is similarly defined. Standard errors are robust and clustered by ja (model * age).
Table 7: Alternative discount rates and time horizon

Dependent variable: Vehicle price

<table>
<thead>
<tr>
<th>Specification</th>
<th>(1)</th>
<th>(2) 5%</th>
<th>(3) 25%</th>
<th>(4) 45%</th>
<th>(5) 3 yr horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G_{jat}$</td>
<td>-0.25</td>
<td>-0.15</td>
<td>-0.36</td>
<td>-0.59</td>
<td>-0.43</td>
</tr>
<tr>
<td>$[-\gamma/\eta]$</td>
<td>(0.11)</td>
<td>(0.07)</td>
<td>(0.16)</td>
<td>(0.25)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>ln(market share)</td>
<td>-4.561</td>
<td>-4.382</td>
<td>-4.649</td>
<td>-4.711</td>
<td>-4.735</td>
</tr>
<tr>
<td>$[-1/\eta]$</td>
<td>(1.588)</td>
<td>(1.513)</td>
<td>(1.625)</td>
<td>(1.650)</td>
<td>(1.659)</td>
</tr>
<tr>
<td>ln(nest share)</td>
<td>410</td>
<td>496</td>
<td>365</td>
<td>329</td>
<td>322</td>
</tr>
<tr>
<td>$[\sigma_1/\eta]$</td>
<td>(581)</td>
<td>(552)</td>
<td>(596)</td>
<td>(606)</td>
<td>(611)</td>
</tr>
<tr>
<td>Observations</td>
<td>546,014</td>
<td>546,014</td>
<td>546,014</td>
<td>546,014</td>
<td>546,014</td>
</tr>
<tr>
<td>ja groups</td>
<td>24,665</td>
<td>24,665</td>
<td>24,665</td>
<td>24,665</td>
<td>24,665</td>
</tr>
</tbody>
</table>

Notes: The sample includes monthly observations from January 2003 to March 2008 of all passenger cars and light trucks age 1-25. See text for calculation of expected gas costs. All specifications include monthly time dummies. Column (1) matches the last column of the Table 3 and shows a nested logit model with vehicle class as the only nest. Here, nest share denotes the share of a vehicle within its class. A 15% annual discount rate is assumed in the calculation of gas costs. Columns (2)-(4) use a 5%, 25%, and 45% discount rate in the calculation of gas costs, respectively. Column (5) uses a 15% discount rate but only accounts for the next 3 years of gas costs. Standard errors are robust and clustered by $ja$ (model * age).
Table 8: Alternate gas price expectations and vehicle usage

<table>
<thead>
<tr>
<th>Specification</th>
<th>(1) Base</th>
<th>(2) See notes</th>
<th>(3) Martingale</th>
<th>(4) Fix VMT</th>
<th>(5) Variable VMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{jat}$</td>
<td>-0.25</td>
<td>-0.27</td>
<td>-0.11</td>
<td>-0.20</td>
<td>-0.21</td>
</tr>
<tr>
<td>$[-\gamma/\eta]$</td>
<td>( 0.11 )</td>
<td>( 0.11 )</td>
<td>( 0.06 )</td>
<td>( 0.12 )</td>
<td>( 0.12 )</td>
</tr>
<tr>
<td>ln(market share)</td>
<td>-4.561</td>
<td>-3.563</td>
<td>-3.362</td>
<td>-2.623</td>
<td>-4.420</td>
</tr>
<tr>
<td>$[-1/\eta]$</td>
<td>( 1.588 )</td>
<td>( 1.185 )</td>
<td>( 1.079 )</td>
<td>( 0.913 )</td>
<td>( 1.548 )</td>
</tr>
<tr>
<td>ln(nest share)</td>
<td>410</td>
<td>444</td>
<td>711</td>
<td>576</td>
<td>425</td>
</tr>
<tr>
<td>$[\sigma_1/\eta]$</td>
<td>( 581 )</td>
<td>( 472 )</td>
<td>( 425 )</td>
<td>( 497 )</td>
<td>( 566 )</td>
</tr>
</tbody>
</table>

| Observations        | 546,014  | 546,014       | 546,014        | 282,507     | 546,014          |
| ja groups           | 24,665   | 24,665        | 24,665         | 10,740      | 24,665           |

Notes: The sample includes monthly observations from January 2003 to March 2008 of all passenger cars and light trucks age 1-25. See text for calculation of expected gas costs. All specifications include monthly time dummies. Column (1) matches the last column of the Table 3 and shows a nested logit model with vehicle class as the only nest. Nest share denotes the share of a vehicle within its class. Column (2) uses oil futures to predict future gasoline prices, but does not include gas costs after available futures data are available. Columns (3) and (4) assume that gas prices are martingale in the calculation of expected gas costs. Column (4) also assumes that each vehicle survives exactly 13 years, and its annual vehicle miles traveled (VMT) is 12,000. Vehicles over age 8 are dropped from specification (4) only. Column (5) assumes that vehicle usage changes with gas prices; see section 6.4 for a complete description. Standard errors are robust and clustered by ja (model * age).

Table 9: Alternate time periods

<table>
<thead>
<tr>
<th>Specification</th>
<th>(1) Base</th>
<th>(2) Jan 03 - Dec 08</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{jat}$</td>
<td>-0.25</td>
<td>-0.51</td>
</tr>
<tr>
<td>$[-\gamma/\eta]$</td>
<td>( 0.11 )</td>
<td>( 0.08 )</td>
</tr>
<tr>
<td>ln(market share)</td>
<td>-4.561</td>
<td>-3.822</td>
</tr>
<tr>
<td>$[-1/\eta]$</td>
<td>( 1.588 )</td>
<td>( 1.468 )</td>
</tr>
<tr>
<td>ln(nest share)</td>
<td>410</td>
<td>993</td>
</tr>
<tr>
<td>$[\sigma_1/\eta]$</td>
<td>( 581 )</td>
<td>( 442 )</td>
</tr>
</tbody>
</table>

| Observations      | 546,014  | 628,554             |
| ja groups         | 24,665   | 25,319              |

Notes: The sample includes monthly observations from January 2003 to March 2008 of all passenger cars and light trucks age 1-25. See text for calculation of expected gas costs. Both specifications include monthly time dummies. Column (1) matches the last column of the Table 3 and shows a nested logit model with vehicle class as the only nest. Here, nest share denotes the share of a vehicle within its class. Column (2) includes monthly data from January 2003 to December 2008. Standard errors are robust and clustered by ja (model * age).