CEEPR Center for Energy and Environmental Policy Research

# Gasoline Prices, Fuel Economy, and the Energy Paradox

by
Hunt Allcott and Nathan Wozny

10-003 March 2010

A Joint Center of the Department of Economics,
MIT Energy Initiative, and Sloan School of Management

# Gasoline Prices, Fuel Economy, and the Energy Paradox

Hunt Allcott and Nathan Wozny

March 29, 2010

**Abstract**

It is often asserted that consumers purchasing automobiles or other goods and services underweight the costs of gasoline or other "add-ons." We test this hypothesis in the US automobile market by examining the effects of time series variation in gasoline price expectations on the prices and market shares of vehicles with different fuel economy ratings. When gas prices rise, demand for high fuel economy vehicles increases, pushing up their relative prices. Market share changes - increased production of high fuel economy vehicles and scrappage of low fuel economy vehicles - attenuate these price changes. Intuitively, the less that equilibrium vehicle prices and shares respond to changes in expected gasoline prices, the less that consumers appear to value gasoline costs.

We estimate a nested logit discrete choice model using a remarkable dataset that includes market shares, characteristics, expected usage, and transaction price microdata for all new and used vehicles available between 1999 and 2008. To address simultaneity bias, we introduce a new instrument for used vehicle market shares, based on the fact that gasoline prices cause variation in new vehicle shares that then persists over time as the vehicles move through resale markets. Our results show that US auto consumers are willing to pay just \$0.61 to reduce expected discounted gas expenditures by \$1. We incorporate the estimated parameters into a new discrete choice approach to behavioral welfare analysis, which suggests with caution that a paternalistic energy efficiency policy could generate welfare gains of \$3.6 billion per year.

**JEL Codes:** D03, L62, Q41.

---

We thank Alberto Abadie, Orley Ashenfelter, David Autor, Lucas Davis, Henry Farber, Kelly Gallagher, Penny Goldberg, Michael Greenstone, Jerry Hausman, Mark Jacobsen, James Kahn, Lutz Kilian, David Laibson, David Lee, Henry Lee, Robin Lee, Lars Lefgren, Alessandro Lizzeri, Erich Muehlegger, Sendhil Mullainathan, Richard Newell, Ariel Pakes, Nancy Rose, Jesse Rothstein, Stephen Ryan, Jim Sallee, Jorge Silva-Risso, Rob Stavins, Chris Timmins, Roger von Haefen, Glen Weyl, Matt White, and seminar participants at Analysis Group, Carnegie Mellon, Colorado School of Mines, Harvard, Kent State University, Mathematica Policy Research, MIT, Princeton, Stanford, University of Delaware, and University of Toronto 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 JD Power for providing vehicle transaction data. Thanks to RL Polk for providing vehicle registration data and to Lonnie Miller and Ray Alvarado at RL Polk for helpful conversations. We thank Manheim Consulting for providing auction data and Gary George, Nirmeet Kacheria, Pete Sauber, and Tom Webb at Manheim for sharing their expertise. We thank Carmen Collyns and Jordon Ricks for research assistance on this and related projects. We especially appreciate the meticulous research assistance of Sounman Hong.

# 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 attentive to ongoing management fees than to upfront payments (Barber, Odean, and Zheng 2005). Shoppers are less elastic to sales taxes than to prices (Chetty, Looney, and Kroft 2009). Consumers' tradeoffs between the purchase price and future energy costs of air conditioners imply relatively high discount rates (Hausman 1979).[1]

Similarly, it is often asserted[2] that vehicles' gasoline costs are not salient to automobile consumers at the time of purchase, and that consumers thus do not fully account for these future costs when choosing between vehicles. As a result, consumers choose lower fuel economy automobiles, with higher resulting fuel expenditures, than they would in their private optimum. 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. The purported undervaluation of gasoline costs would also help explain what Jaffe and Stavins (1994) call the "Energy Paradox": that consumers and firms have been remarkably slow to adopt apparently high-return energy efficient technologies.[3]

Externalities related to national security and climate change would exacerbate the private welfare losses from consumers' potential undervaluation of fuel economy. There has been substantial debate over whether these externalities should be internalized through gasoline taxes or Corporate

[1] Hausman estimates that consumers implicitly use a discount rate of 25 percent per year when they trade of purchase prices and future energy costs of new air conditioners. He concludes (p. 51), "Yet this finding of a high individual discount rate does not surprise most economists. At least since Pigou, many economists have commented on a "defective telescopic faculty." A simple fact emerges that in making decisions which involve discounting over time, individuals behave in a manner which implies a much higher discount rate than can be explained in terms of the opportunity cost of funds available in credit markets. Since this individual discount rate substantially exceeds the social discount rate used in benefit-cost calculations, the divergence might be narrowed by policies which lead to purchases of more energy-efficient equipment."

[2] See, for example, Greene (1998) and Parry, Walls, and Harrington (2007).

[3] Various explanations have been proposed for this apparent anomaly, including imperfect information, credit constraints, principal-agent problems (Murtishaw and Sathaye 2006), some form of bounded rationality (DeCanio 1993), and that discount rates do not properly model hysteresis and irreversible investment under uncertainty (Hassett and Metcalf 1993). Yet another explanation is that there may be no "Paradox" at all: analysts' estimates of the returns to investments that improve energy efficiency may be overly optimistic (Metcalf and Hassett 1999).

Average Fuel Economy (CAFE) standards (e.g. Bento, *et al,* 2009). Economists often 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. CAFE standards, by contrast, only bind on the extensive margin.[4] If consumers undervalue future fuel costs when they choose between vehicles, however, their extensive margin response to gasoline taxes would not be optimal, and if the undervaluation is sufficient, CAFE standards might be preferred. Indeed, one of the leading economic arguments for CAFE and other energy efficiency standards is that they could increase welfare by forcing consumers to own more energy efficient durable goods, regardless of whether their choices indicate that they want them.[5]

A central problem in taking the paternalistic stance on fuel economy standards is the dearth of evidence on whether automobile consumers actually are or are not misoptimizing. The rational model provides our null hypothesis: that consumers are willing to pay one extra dollar in vehicle purchase price to decrease the expected present value of future gasoline costs by one dollar. Although many phrases could be used, for expositional purposes we will say that rejecting this hypothesis is evidence that consumers "misvalue gasoline costs."[6] This paper tests the null hypothesis using extraordinary micro- and market-level data on the prices, quantities, characteristics, and usage of all passenger vehicles in the United States between 1999 and 2008.

Our empirical test is based on the intuition that the increase in gasoline prices over the past decade should increase the relative prices and market shares of high- vs. low-fuel economy vehicles. Indeed, media reports and academic analyses have documented that as gasoline prices rise, the market shares of new high fuel economy vehicles rise (Klier and Linn 2008), the scrappage of used low fuel economy vehicles increases (Li, Timmins, and Von Haefen 2009), and the relative prices of both new and used vehicles with low fuel economy drop (Busse, Knittel, and Zettelmeyer 2009). The above null hypothesis, however, does more than predict that gasoline prices should affect

[4] The higher fuel economy vehicles required under CAFE require less fuel to operate per mile, and thus consumers actually have the incentive on the intensive margin to increase driving. This is often called the "rebound effect."

[5] This "paternalistic" argument for fuel economy standards is discussed in the government's Regulatory Impact Analysis of the new CAFE standards (NHTSA 2009, page 335) and is suggested by Fischer, Harrington, and Parry (2007), Greene (1998), Greene, Patterson, Singh, and Li (2005), and Parry, Walls, and Harrington (2007), among others. Some of these analyses do not necessarily advocate the position given the lack of empirical support for misoptimization. Hausman and Joskow (1982) discuss this argument in the context of appliance energy efficiency standards.

[6] Related research might use different terms to describe undervaluation of gasoline costs, such as myopia, misoptimization, inattention, shrouding, salience, high implicit discount rates, and naivete.

relative demand for vehicles of different fuel economy ratings: it predicts *how much* demand should be affected. Finding that changes in relative prices and shares are smaller than predicted suggests that consumers undervalue gasoline costs and fuel economy when they purchase vehicles.

In principle, consumers' valuation of fuel economy and future gasoline costs could be estimated in a hedonic or discrete choice framework using variation in the prices and fuel economy ratings in one cross section of vehicles. Since energy costs are a function of the product's energy efficiency and of energy prices, however, this valuation can alternatively be estimated from time-series changes in energy price expectations. We adopt this approach, using panel data on vehicle markets with vehicle-specific fixed effects. This allows our estimator to be unbiased even if a vehicle's fuel economy is correlated with its unobserved characteristics.

As proposed by Kahn (1986)[7], the panel approach is simplest if one assumes that neither new vehicle supply nor used vehicle scrappage rates respond to gas prices. In this intuitive model, the relative price of a used vehicle should decrease by one dollar for each one-dollar increase in the relative present discounted value of expected future gasoline costs. Unfortunately, the observed response of market shares to gas price changes biases that approach towards concluding that consumers undervalue gasoline costs. Higher gas costs lead to a decrease in production of low fuel economy new vehicles. The reduced supply in turn increases the prices of both these new vehicles and used vehicles that are good substitutes. Similarly, production of high fuel economy vehicles would go up, reducing the prices of new and used high MPG vehicles. The responsiveness of used vehicle prices to gas costs would therefore be smaller than would be expected if quantities were assumed to be fixed. This responsiveness is further attenuated if scrappage of low fuel economy vehicles increases with gas prices.

We account for vehicle quantities and substitution patterns using a discrete choice model of vehicle demand, where consumers' utility from owning a vehicle is allowed to depend separately on discounted future gasoline costs and the purchase price. To account for unobserved heterogeneity in consumers' preferences, we use a nested logit model. The benefit of the nested logit specification

[7] Kilian and Sims (2006) and Sallee, West, and Fan (2009) build on Kahn's (1986) fundamental approach. Other work that examines how vehicle prices adjust in response to gasoline prices include Sawhill (2008), Langer and Miller (2009), and Austin (2008). Verboven (1999) estimates the discount rates implied by differences between the prices of gasoline and diesel vehicles in Europe. Ohta and Griliches (1986) examine whether the 1970s gasoline price shocks affected consumers' valuations of vehicle characteristics.

is that it gives a simple market-level relationship between equilibrium vehicle prices, market shares, and gasoline costs while parsimoniously modeling substitution patterns across similar vehicles.

In estimating the demand equation that results from our model, we face the standard simultaneity problem that vehicle market shares (and prices) may be correlated with unobserved vehicle characteristics. Accordingly, we instrument for market shares by exploiting the fact that the demand for a new vehicle with low fuel economy is higher in years when gasoline prices are low. At any time in the future, the quantity available of the (now used) vehicle produced in that year will therefore be higher than the quantity of the same model produced in a model year when gasoline prices were high. Crucially, this within-model variation in quantity of used vehicles should be independent of unobserved product attributes. Our paper therefore introduces into the literature a new instrument for automobile market shares or prices, based on the interaction of fuel economy with the gasoline price in the model year in which the vehicle was produced. This may prove to be more broadly useful as an alternative to the standard Berry, Levinsohn, and Pakes (1995) instrumental variables procedure.

We estimate the model with perhaps the largest collection of data ever used in the economics literature on the automobile industry. From microdata on 57 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 comprehensive vehicle registration data, we observe the national-level 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 the 25,000-household National Household Travel Survey, covering vehicle ownership and vehicle-miles traveled.

We formulate our assumptions to conservatively bias us against finding that consumers undervalue gasoline costs. We find, however, undervaluation for any plausible set of assumptions about gasoline cost expectations, vehicle survival probabilities, vehicle-miles traveled, and other parameters. We conservatively estimate that between 1999 and March 2008, American auto consumers were willing to pay only sixty-one cents to reduce expected discounted gas expenditures by one dollar.

Under the assumption that our empirical results are driven by inattention to future gasoline costs, we then compute the welfare implications of misoptimization. To carry this out, we apply the framework of Bernheim and Rangel (2009) and related analyses to introduce a new and highly tractable approach to behavioral welfare analysis in a discrete choice setting. We analyze a counterfactual "Behavioral Feebate" policy that imposes sales taxes that increase in a vehicle's expected future gasoline consumption by an amount such that consumers purchase their privately-optimal vehicles. Given our parameter estimates and stylized modeling assumptions, the welfare gains from such a policy are $15 per potential vehicle owner per year. Across approximately 240 million potential vehicle owners, this sums to $3.6 billion annually.

These findings have implications in several domains. First, a cap-and-trade program to internalize the marginal damages of carbon dioxide emissions would act on the automobile market through an increase in gasoline prices. For a cap-and-trade or a comparable Pigouvian tax to achieve the first best requires that all consumers arrive at their own private optima given the new higher relative prices of pollution-intensive goods. If automobile buyers undervalue future gasoline prices, other sectors will have to abate more carbon to satisfy a carbon emissions cap, and the marginal cost of abatement will be above the optimum and will not be equal across sectors. Through this logic, our finding adds empirical justification for extending the discussion of tax salience (e.g. Finkelstein (2009) and Chetty, Looney, and Kroft (2009)) into "environmental tax salience."

Second, 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. Analyses including Bento, *et al*, (2009), Berry, Levinsohn, and Pakes (1995, 2004), Goldberg (1995, 1998), Jacobsen (2008), and Nevo (2002) all use an estimate of consumers' demand for higher fuel economy vehicles. Our analysis is complementary to this body of work in that it provides a careful estimate of an essential demand parameter.

Third, evidence that consumers are inattentive to future product costs also has important implications for how firms behave in equilibrium. "Myopic" or "unsophisticated" consumers, in the sense of Gabaix and Laibson (2006) and Ellison (2005), may be one reason why firms set low markups on base products such as credit card interest rates, razors, and printers and set high markups for add-ons such as late fees, razor blades, and printer cartridges. Although automobile manufacturing firms do not sell gasoline (the "add-on"), a related model can be applied in this industry, as the fuel economy embodied in a vehicle determines gasoline demand, and improving fuel economy is costly. Furthermore, if gasoline costs are in essence a "shrouded attribute" in consumers' decisions, this reduces manufacturers' ability to exploit economies of scale in producing high fuel economy vehicles and dulls their incentives to direct technological change toward reducing the cost of such vehicles. This suggests additional channels through which regulations such as Corporate Average Fuel Economy standards and fuel economy information labels can affect consumer welfare.

The paper progresses as follows. In section 2, we provide an overview of how we conceptualize this problem, making the connection between features of the economic problem and econometric identification. 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 that we have gathered and devotes particular attention to the construction of a vehicle's discounted expected future gasoline costs. Section 6 presents our main results and an extensive series of robustness checks. Section 7 presents the theory and results of the welfare calculation.. Section 8 concludes with a note of caution on whether this analysis should be used to advocate for paternalistic energy efficiency policies.

## 2 Conceptualizing the Problem

All else equal, an optimizing consumer should be willing to pay \$1 more for a product that entails \$1 less in discounted future costs. The fundamental goal of this paper is to test whether observed automobile market equilibria are consistent with this condition. Our test requires us to construct a framework that predicts how gasoline price-induced demand shifts affect equilibrium vehicle prices and quantities, and then to compare the predictions to data. This section introduces the economic intuition for our approach; our formal model is introduced in section 3. One key takeaway from this section will be that both vehicle prices and market shares must be endogenized - a simpler estimator that holds quantities fixed would generate biased estimates.

A number of analyses have attempted to measure the importance of energy efficiency in the choice of energy using durable goods, in a long and active literature on "implicit discount rates."

The most common identification strategy has been to exploit variation in the prices and energy efficiencies in a cross-section of products. Cross-sectional identification strategies were used in 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 discrete choice framework, conditional on other product characteristics, a one dollar increase in purchase price should be associated with the same decrease in market share as a one dollar increase in lifetime energy costs. In a hedonic regression, a one dollar increase in energy costs should be associated with a one dollar decrease in price.

For such an estimator to be unbiased, any unobserved characteristics must be uncorrelated with energy efficiency, and the functional form of any observed and correlated characteristics must be correctly specified. With automobiles, this assumption is likely to be problematic. Fuel economy is highly 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.[8] Furthermore, fuel economy is affected by styling decisions that 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 observe or quantify. In a cross section, fuel economy is negatively correlated with price, which suggests that low fuel economy vehicles may have more unobserved characteristics that increase utility.

The ability to look "within" the same vehicle over time as gasoline prices change obviates the need to make assumptions about how the vehicle's unobserved characteristics are correlated with fuel economy. As a concrete example, consider comparing the price of a model year 2001 Honda Civic in 2006 to a model year 2002 Honda Civic in 2007. If all else were equal other than gasoline prices, the change in price of this 5 year old Civic from 2006 to 2007 would be the response to the change in expected gas costs between the two years. An increase in gasoline prices should increase its purchase price relative to a lower-fuel economy vehicle, and decrease its purchase price relative to a higher-fuel economy vehicle. What we must do now is think through *how much* these relative prices should change.

[8] 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 estimation of automobile demand in characteristic space sometimes gives the "wrong" sign on fuel economy.

We add some mathematical structure to provide intuition for the problem. The simple model that follows will motivate the need for the formal discrete choice model that will be presented beginning in the subsequent section. Consider a very simple world where consumers choose between a vehicle and some outside option. A demand function is:

$$q = \alpha - \eta p - \gamma G \tag{1}$$

In this equation, $q$ is the vehicle's quantity, $\alpha$ is an intercept, $p$ is the vehicle's purchase price, and $G$ is the discounted gasoline costs over the vehicle's lifetime, which we assume to be the same for all consumers. The variable $G$ will depend on a discount rate, future gas prices, fuel economy, and the vehicle's usage and scrappage probability over time, and we will later return to these issues in great detail. We allow consumers to value purchase price and gas costs unequally. We can also re-arrange this equation to get equilibrium price on the left-hand side:

$$p = \frac{1}{\eta}(-\gamma G + \alpha - q) \tag{2}$$

Assume for the next few paragraphs that the vehicle's quantity $q$ is constant. Equation (3) shows that if $\gamma = \eta$, as in our null hypothesis, a one dollar increase in $G$ would cause a one dollar decrease in $p$. Intuitively, if quantity is constant, the same consumer, with the same willingness to pay, sets the price. To keep this consumer indifferent between the vehicle and the most attractive outside option, the overall product cost $p + G$ must stay the same as $G$ changes.

The ideal dataset would allow us to observe the same consumers, with the same choice set of vehicles, in the same fixed market shares as gas prices change over time. Under these conditions, the marginal consumer for each vehicle stays the same. The changes in gasoline prices would generate variation in $G$ - smaller or larger changes, depending on the fuel economy of each vehicle. In this ideal world, $\frac{\gamma}{\eta}$ could then be consistently estimated using the following panel regression:

$$p_{jat} = -\frac{\gamma}{\eta} G_{jat} + \psi_{ja} + \tau_t + \varepsilon_{jat} \tag{3}$$

In this equation, $\psi_{ja}$ is a constant for each "vehicle" of model $j$ of age $a$, equal to a marginal consumer's willingness to pay to operate the vehicle. The variable $\tau_t$ captures changes in the overall average price level at time $t$, and $\varepsilon_{jat}$ is some econometric error. This approach is qualitatively consistent with the analogous specifications in Kahn (1986) and the literature that follows his approach.[9]

Now let us relax the assumption the the vehicle's quantity is constant. Either through new vehicle sales or scrappage of used vehicles, we now recognize that the market share of the vehicle in equation (1) is not fully inelastic. The quantity supplied $q^S$ is:

$$q^S = \beta_0 + \beta_1 p \tag{4}$$

Equating quantity supplied with quantity demanded, the the equilibrium vehicle price is:

$$p = -\frac{\gamma}{\eta + \beta_1} G + \frac{\alpha_0 - \beta_0}{\eta + \beta_1} \tag{5}$$

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 (3) when supply is not fully inelastic would result in a downward bias in an estimate of $\frac{\gamma}{\eta}$, which would lead the analyst to incorrectly conclude that consumers undervalue gasoline costs. In other words, if quantity $q$ is correlated with $G$, then omitting $q$ in the estimation of (3) will bias the estimated coefficient on $G$.

Production of new vehicles is much more price elastic than scrappage of used vehicles. Given this, one response to the problem of elastic supply would be to analyze only the used vehicle market, assuming that the effect of gas price expectations on scrappage rates is negligible. However, consumers also substitute between new and used vehicles with similar characteristics. If an increase in gas prices leads to a decrease in production of new low fuel economy vehicles, this will increase

[9] 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 adjusted to gasoline price changes in the 1970s.

the willingness to pay for a substitutable used low fuel economy vehicle. Again referring to equation (2), this substitution will generate a positive correlation between the $G$ and the demand intercept $\alpha$. Not accounting for this generates another correlation between $G$ and the error term in estimating equation (3), which would further bias downward an estimate of $\frac{\gamma}{\eta}$ and bias the analysis toward concluding that consumers undervalue gasoline costs.

# 3 Model

In this section, we describe our discrete choice model, which addresses the concerns from the previous section by endogenizing substitution patterns and changes in market shares. This model will be a modification of the standard framework in the industrial organization discrete choice literature, e.g. Berry (1994). In our static discrete choice model, consumers derive utility from owning a vehicle and from consuming a numeraire 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 $\Omega$.

In each period, consumers choose from a set of new and used models $j = 1, ..., J_t$, 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. As in all static discrete choice models, consumers choose a vehicle in each period and expect to hold the vehicle for the remainder of its life.[10] Consumer $i$ receives indirect utility $u_{ijat}$ from purchasing vehicle $ja$ in year $t$:

$$u_{ijat} = \eta(w - p_{jat}) - \gamma G_{jat} + \psi_{jat} + \epsilon_{ijat} \tag{6}$$

In this equation, $w$ is the consumer's wealth, $p_{jat}$ is the purchase price, and $G_{jat}$ is the discounted present value of future gasoline costs over the vehicle's lifetime. If consumers value purchase price and future 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 the construction of this variable

[10] In reality, the consumer's true problem is dynamic: at any point in time, she has the opportunity to re-sell the vehicle and purchase a new one. In the results section and in Appendix A.1, we return to the assumptions required to simplify the consumer's demand problem from a dynamic choice to a static choice.

in great detail in section 5. Note that we assume that consumers are risk neutral,[11] and $G_{jat}$ will be constructed using expectations. The variable $\psi_{jat}$ 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.

As described in section 2, it is important to capture how a vehicle's price might be affected by changes in the prices of substitutes: the model must capture, for example, how a decrease in the price of new SUVs should affect demand for used SUVs. This requires a reasonable model of how the individual's unobserved "taste shock" $\epsilon_{ijat}$ varies across vehicles. If taste shocks are uncorrelated across vehicles, as in the homogeneous consumer logit model, substitution is proportional to market shares. For example, as the price of a new SUV increases, the homogeneous consumer model predicts that consumers would substitute equally to a used compact car and a used SUV that had the same market share. In reality, we should expect more substitution to the used SUV, given that it is a more similar product. Consumers that transport large families in an SUV will have a hard time substituting to compact cars.

More realistic substitution patterns are captured econometrically by parameterizing correlations in unobserved tastes $\epsilon_{ijat}$ across vehicles. We use a nested logit framework, which allows consumer's idiosyncratic preferences to be correlated across vehicles within the same predetermined set of vehicles, or "nest": $corr(\epsilon_{ijat}, \epsilon_{ij'a't})$ is nonnegative when $ja$ and $j'a'$ are in the same nest and zero otherwise.[12] We will estimate a parameter $\sigma$ related to these within-nest correlations.[13]

[11] Risk aversion does not bias our estimator if uncertainty over future gasoline prices is constant, or more weakly, uncorrelated with the level of gas prices. Constant expected volatility in future gasoline prices would result in a risk averse consumer having a lower willingness to pay for any particular vehicle. Because lower MPG magnifies the effect of gasoline prices on a consumer's gasoline expenditures, this effect would be larger for low MPG vehicles. However, any discount in the vehicle's price due to constant gas price volatility predicted by a model of risk averse consumers would be absorbed in the fixed effect.

In reality, however, implied volatility was positively correlated with gasoline prices over the study period. If consumers are risk averse, this should cause consumers to want to substitute away from low MPG vehicles as volatilities (and also prices) rose. Since we find that relative prices of low fuel economy vehicles did not fall as much as predicted when gasoline prices rise, the risk neutrality assumption actually strengthens our qualitative result.

[12] Another common way of parameterizing unobserved taste shocks is through a random coefficients model, which can allow preferences for continuous attributes such as horsepower and weight to vary across the population. We use the nested logit because our choice set is unusually large and because it produces a transparent, simple log-linear relationship between market-level prices and shares.

[13] In particular, the cumulative distribution function for $\epsilon_{ijat}$ for all $ja$ for individual $i$ at time $t$ is: $F(\cdot) = \exp\left[-\sum_{n\in\mathcal{N}}\left(\sum_{ja\in\mathcal{B}_n} e^{-\epsilon_{ijat}/(1-\sigma)}\right)^{1-\sigma}\right]$. $\mathcal{N}$ is the set of all nests of vehicles, and $\mathcal{B}_n$ is the set of vehicles in nest $k$. $\sigma$ is a parameter related to the within-nest correlation of utilities and will be estimated in the model. As $\sigma$ approaches one, the within-nest correlation of utilities approaches one. If $\sigma = 0$, the standard logit model is recovered. This distribution can be extended to accommodate multiple nests or separate $\sigma$ parameters for each nest.

As in other nested logit models, the nests are specified ex ante and determine the structure of substitution patterns allowed by the model. Nests must be 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: a consumer is 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' ability to substitute among vehicles of different sizes and therefore different fuel economy ratings. This, in turn, would lead us to overstate consumers' responsiveness to changes in gas price expectations.

It is well-known (e.g. Berry 1994) that if the utility of the outside good is normalized to zero, the nested logit choice probabilities can be aggregated over the population to give a market-level relationship between prices and shares:

$$\ln s_{jat} - \ln s_{0t} = -\eta p_{jat} - \gamma G_{jat} + \sigma \ln(s_{jat}/s_{nt}) + \psi_{jat} \tag{7}$$

In this equation, $s_{jat}$ is the market share of vehicle $ja$, $s_{0t}$ is the share of the outside option, and $s_{nt}$ is the combined market share of all vehicles in nest $n$, of which vehicle $ja$ is a member.

Recall that the purpose of the model is to test whether $\eta = \gamma$, i.e. whether consumers are indifferent between one dollar in purchase price and one dollar in future gasoline costs. The market-level relationship between equilibrium prices and quantities implied by this discrete choice framework will be the basis of our empirical test. Finding from market data that $\gamma < \eta$ would suggest that consumers underweight future gasoline costs relative to purchase price in their decision.[14]

# 4 Empirical Strategy

Any consistent estimator of equation (7) must address two problems. The first is simultaneity bias: prices and quantities are affected by unobserved product attributes that enter utility func-

[14] We assume that $\eta$ and $\gamma$ are constant and homogeneous in the population, which produces this simple hypothesis that can be tested with a linear model. The marginal utility of money is likely to vary across consumers, which explains differences in preferences for luxury vehicles and vehicles of different ages. We proxy for this heterogeneity using nests for vehicle age and luxury vehicles, as described in detail in section 7.

tions through $\psi$. The second is the potential correlation of fuel economy and unobserved product attributes in the cross section.

To address the first problem, simultaneity bias, we will in the following subsection introduce instrumental variables. These instruments will be more intuitive as instruments for market shares instead of prices, so purely for intuition and notational convenience, we rearrange the market-level relationship in equation (7) so that price is on the left hand side:

$$p_{jat} = -\frac{1}{\eta}\left(\ln s_{jat} - \ln s_{0t}\right) - \frac{\gamma}{\eta}G_{jat} + \frac{\sigma}{\eta}\ln(s_{jat}/s_{nt}) + \frac{\psi_{jat}}{\eta} \tag{8}$$

The second problem is that average utility obtained from a vehicle $\psi_{jat}$ depends on average preferences for observed and unobserved characteristics. It is theoretically possible to estimate equation (8) using a cross section of vehicles with different prices and fuel economy ratings. This would require, however, that we observe and parameterize vehicle characteristics well enough to assume that no unobserved part of $\psi_{jat}$ is correlated with fuel economy.

Instead, our panel identification strategy exploits model-by-age fixed effects $\psi_{ja}$. Since observable vehicle characteristics are effectively identical across the years of vehicle $ja$, the deviation from vehicle average utility is a year-specific unobservable: $\psi_{jat} = \psi_{ja} + \xi_{jat}$. After adding model year fixed effects $\mu_{t-a}$ and a time dummy $\tau_t$ to absorb the outside option share $\ln s_{0t}$ any overall shift in the price level of vehicles, equation (8) becomes:

$$p_{jat} = -\frac{\gamma}{\eta}G_{jat} - \frac{1}{\eta}\ln s_{jat} + \frac{\sigma}{\eta}\ln(s_{jat}/s_{nt}) + \frac{\psi_{ja}}{\eta} + \tau_t + \mu_{t-a} + \frac{\xi_{jat}}{\eta} \tag{9}$$

Equation (9) resembles the reduced form equation (3) from section 2. Were we willing to assume that market shares are fixed, or more weakly, uncorrelated with $G$, then we could leave them in the error term. $\frac{\gamma}{\eta}$ could then be identified as the coefficient of $G_{jat}$ in an estimation of equation (3). Because of the evidence that both new vehicle sales and used vehicle scrappage respond to gasoline prices, however, this estimator would be biased.

Our specification requires $E[\xi G] = 0$, but allows $E[\psi G] \neq 0$. In words, the model-year specific unobservable characteristic must be uncorrelated with fuel economy and gas prices, but the fixed effects allow vehicle characteristics that are fixed within models across model years to be correlated with fuel economy. Even after using fixed effects, however, the model year-specific unobservable characteristic $\xi_{jat}$ could still be correlated with market shares if, for example, a feature that is specific to particular model year affects both price and share: $E[\xi s] \neq 0$. To address this, we need an instrument that generates variation in market shares that is uncorrelated with unobserved quality. As we will see momentarily, the fact that new vehicle sales respond to gasoline prices suggests an instrument for used vehicle quantities. Interestingly, the problem that motivated this approach - that market shares respond to gasoline prices - is also part of the solution.

### 4.1 Instruments and Two-Stage Least Squares

Our instrument exploits the stylized fact that 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. For example, the increase in gasoline prices from 2004 to 2005 means that there should be more two-year old gas guzzlers on the road in 2006 compared to 2007.

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 $G_{j0(t-a)}$. This instrument acts conditional on model year dummy variables, meaning that vehicles that have high values of $G_{j0(t-a)}$ relative to other vehicles produced in the same year are expected to have lower sales. For vehicles of different model years within the model-by-age fixed effect groups, this generates variation in market shares that is independent of the demand shifter $\xi_{jat}$.

We assume that $\ln s_{nt}$, the log market share of nest $n$, is independent of $\xi_{jat}$. Although it is somewhat awkward, this is necessitated by the fact that, because the class-level nests include vehicles of different model years, our instrument does not generate substantial variation in the nest market shares. In practice, the nests include many vehicles, and the share of vehicle $ja$ is a small fraction of the share of nest $n$.[15]

[15] Any bias generated in $\widehat{\frac{\gamma}{\eta}}$ depends on the covariance of $G_{jat}$ and $\ln s_{nt}$, and the covariance of $\ln s_{nt}$ and $\xi_{jat}$. The expected supply response suggests that $G_{jat}$ is likely to be negatively correlated to $\ln s_{nt}$. $-\widehat{\frac{\gamma}{\eta}}$ will be biased

The first stage equation of the two stage least squares regression is:

$$\ln s_{jat} = \alpha_{11} G_{j0(t-a)} + \alpha_{12} G_{jat} + \alpha_{13} \ln s_{nt} + \psi'_{ja} + \tau'_t + \mu'_{t-a} + \xi'_{jat} \tag{10}$$

In this equation, the primes on $\psi'_{ja}$ , $\tau'_t$, $+\mu'_{t-a}$, and $\xi'_{jat}$ indicate that the concept of the variable is the same as in the second stage, but the estimated value may of course be different in the first stage. The second stage is:

$$p_{jat} = -\frac{\gamma}{\eta} G_{jat} - \frac{1-\sigma}{\eta} \widehat{\ln s_{jat}} - \frac{\sigma}{\eta} \ln s_{nt} + \frac{\psi_{ja}}{\eta} + \tau_t + \mu_{t-a} + \frac{\xi_{jat}}{\eta} \tag{11}$$

# 5 Data

We have assembled from multiple sources a comprehensive dataset of the average prices, quantities, and characteristics of all passenger vehicle models registered in the US, in monthly cross sections from January 1999 to December 2008. Our dataset comprises 1.1 million observations. Table 1 presents descriptive statistics. Appendix 2 provides extensive additional detail on data sources and variable construction.

Used vehicle prices are based on auction data obtained from Manheim, the largest automobile auctioneer in the United States. The principal buyers in the auctions are dealers who then resell the used vehicles to customers. We have data on each of the approximately 4 million vehicle transactions 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 $t$, adjusting for the vehicle's condition, odometer reading, and region and method of sale. While only about one in four used vehicles traded passes through an auction (Manheim 2009), the auction market is the largest source of transaction price data. Furthermore, the Kelley Blue Book and other

upward (towards zero) if $\ln s_{nt}$ is positively correlated with $\xi_{jat}$. This occurs to the extent that unexplained shifts in equilibrium prices and nest share are driven by shifts in the demand curve rather than shifts in the supply curve. Anecdotal evidence suggests that the automobile market is substantially driven by supply shifts, such as variation in off-lease and rental vehicles entering the auction market. Therefore, while a bias is possible, we expect it to be smaller in magnitude than in a model that does not account for vehicle substitutes.

price guides, which are the starting point for price negotiations in many used vehicle transactions, are largely based on auction prices.

New vehicle prices are from the Power Information Network, a network of dealerships managed by JD Power and Associates. These dealerships report 2 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. We also incorporate used vehicle retail transaction prices from JD Power in specification checks.

We observe national-level quantities in use of each vehicle model in July of each year from 1999 through 2008. These data are from the National Vehicle Population Profile, which we obtained from the automotive market research firm R.L. Polk. The quantities represent all vehicles registered as of July 1, including both 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.

Fuel economy data were obtained 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. 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.

We defined a "vehicle" (in our notation, a $ja$ combination) to capture all possible variation in fuel economy ratings. This entailed disaggregation to the level of vehicle make, model name, trim level, and the number of engine cylinders. The average make and model name combination in our dataset includes four "models"; for example, there are eight different configurations of cars called the model year 2004 Honda Civic (Dx, Vp (Coupe), Hybrid, etc.) that appear as separate "vehicles" in our dataset. Data on new and used vehicle prices and registered quantities are matched using digits from the Vehicle Identification Number that are common within a model and model year.

## 5.1 Expected Discounted Gasoline Costs

We now describe the formulation of the different components of expected discounted gasoline costs $G_{jat}$, including vehicle-miles traveled, survival probability, discount rates, and expected gasoline costs. Given that our parameter of interest $\frac{\gamma}{\eta}$ will be the coefficient on this variable, its construction is especially important for producing convincing results. For example, using a lower discount rate than consumers actually face would inflate $G_{jat}$, thereby biasing $\widehat{\frac{\gamma}{\eta}}$ downward. Alternatively, understating the expected lifetime or usage of the vehicle would deflate $G_{jat}$, biasing $\widehat{\frac{\gamma}{\eta}}$ upward.

Although we use the best available data to construct the components of $G_{jat}$, any of these calculations could be subject to debate. In determining each component, we therefore choose "conservatively," meaning that if $\widehat{\frac{\gamma}{\eta}}$ is biased, it is biased upwards. By erring in this direction, we will show that $\frac{\gamma}{\eta} < 1$ for any plausible set of definitions of the components of $G_{jat}$. Readers interested in even more detail on the construction of $G_{jat}$ should consult the Data Appendix.

The variable $G_{jat}$ is the net present value of expected lifetime gasoline costs over future years $s$:

$$G_{jat} = E_t\left[\sum_{s=t+1}^{t+(L-1-a)} \beta^{s-t} \cdot g_s \cdot m_{ja} \cdot f_{jas}^{-1} \cdot \phi_{jas}\right]$$
$$= \sum_{s=t+1}^{t+(L-1-a)} \beta^{s-t} \cdot E_t[g_s] \cdot E_t[m_{ja}] \cdot E_t\left[f_{jas}^{-1}\right] \cdot E_t\left[\phi_{jas}\right] \quad (12)$$

$L$ denotes the maximum possible lifetime of a vehicle, which we take to be 25 years. The variable $g_s$ is a gasoline price in year $s$, $m_{ja}$ is expected vehicle miles traveled (VMT), $f_{jas}$ 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, and $\beta$ is an annual discount factor.[16] We assume that $G_{jat}$ is homogeneous for all consumers that choose vehicle $ja$ at time $t$.[17]

[16] We assume that vehicles survive with probability one throughout each year, then a fraction determined by $\phi_{jas}$ are removed from the market at the end of the year. We also model that all gasoline costs flow halfway through the year.

[17] In reality, there is substantial variation in vehicle-miles traveled across consumers that own the same vehicle. Similarly, differences across consumers in the proportion of city versus highway driving, other driving behavior, and

The second line of equation (12) 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. A key concern about this assumption is that vehicle miles traveled $m$ are likely to respond to gasoline prices $g$. Indeed, there are a large number of papers that estimate the short run elasticity of gasoline demand (e.g. Hughes, Knittel, and Sperling (2007), Small and Van Dender (2007), and Davis and Kilian (2009)), which often find that the parameter is small but statistically non-zero. If owners of low fuel economy vehicles respond to high gas prices by driving less, then we overstate their gas costs. However, the consumer's utility from vehicle use also decreases in response to the reduced driving. As discussed in Kahn (1986), the Envelope Theorem implies that these changes are equal and opposite to first order, and in our primary specifications, we assume away these effects.

Our primary specification adopts this argument, using VMT predicted from year-2001 gasoline prices and assuming that $\frac{\partial m}{\partial g} = 0$. We also derive an alternative specification that captures the effects of intensive margin elasticity; see Appendix 2 for more detail. As we will later show, this adjustment does not substantially affect the results.

#### 5.1.1 Vehicle-Miles Traveled and Survival Probability

To estimate Vehicle-Miles Traveled (VMT), we use publicly-available data in the 2001 National Household Travel Survey. This is a nationally-representative survey of approximately 25,000 households that report, among many other variables, the age, fuel economy, and vehicle class for each of their 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. These two readings were then used to estimate annualized VMT. We regress annualized VMT on the vehicle's age, class, and fuel

vehicle maintenance generate differences in realized fuel economy. As gasoline price expectations change, the change in relative prices between two vehicles is determined by some marginal consumer who is indifferent between them. In practice, we use the mean VMT for that vehicle since we cannot identify the marginal consumer, but we do not have a reason to believe that this choice generates a systematic bias in the computation of $G_{jat}$.

Note that changes in gasoline prices should lead to a re-sorting of vehicles across consumers with different VMT. As gas prices increase, consumers with relatively high VMT are more likely to switch to a vehicle with higher MPG, but consumers with relatively low VMT would switch to a vehicle with lower MPG. Intuitively, this re-sorting does not affect equilibrium relative prices as long as our the mean VMT for a vehicle is a good proxy for the VMT of a consumer who remains on the margin for that vehicle as gas prices change.

economy, and use these estimates to predict $m_{ja}$ for all vehicles in our sample.

We compute survival probability based on observed survival probabilities in our registration data. As with VMT, we assume that these survival probabilities do not depend on expected gasoline prices. Note that since we regression-adjust vehicle VMT in the individual auction transactions to a standardized value, the left-hand-side variable $p_{jat}$ is not affected by changes in VMT, and thus remaining lifetime, that may be driven by gas price differences over the study period.

#### 5.1.2 Discount Rates

The discount rate $r = \beta^{-1} - 1$ that consumers apply to gasoline costs should reflect the interest rate on the marginal dollar spent on the vehicle. We present two potential benchmarks. First, for a consumer who finances her vehicle, this should be the automobile loan interest rate. The JD Power transaction data include the loan annual percentage rate for vehicles that were financed at the dealership. The transaction-weighted average real interest rate over the study period is 4.7 percent for new vehicles and 9.0 percent for used vehicles.

Second, for a consumer who purchases the vehicle with cash, the opportunity cost could be the expected market returns on an alternative investment. The average real return on the S&P 500 from 1945-2008 was 5.78 percent, but given that the market and oil prices have a very small covariance, the risk-adjusted discount rate for gasoline costs would in fact be close to the risk-free rate.[18] Our primary specification uses an annual real discount rate of 9 percent, which is conservatively at the upper end of these benchmarks.

#### 5.1.3 Gasoline Price Expectations

A perfect measure of consumers' gasoline price expectations $E_t[g_s]$ is not observed. Just as we could frame the analysis as solving for an "implicit discount rate" at which $\widehat{\frac{\gamma}{\eta}} = 1$, we could also solve for

[18] A risk averse consumer with declining marginal utility of consumption would want to risk-adjust returns for their covariance with the market. The Capital Asset Pricing Model (CAPM) allows us to compute the risk-adjusted rate of return that the consumer would require for gasoline purchases, which can also be thought of as disinvestments in gasoline. Annual data from 1945 to 2008 show that oil prices (and therefore gasoline prices) are slightly negatively correlated with market returns, as measured by the S&P 500 stock index. Therefore, the CAPM predicts that a consumer should expect a rate of return on a disinvestment in gasoline that is slightly higher than the risk-free rate of return - by our calculations, about 1.6 percent. The CAPM therefore suggests that our higher discount rate is quite conservative.

any number of formulations of gasoline price expectations which produce that result. The objective of our primary specification is to estimate $\frac{\gamma}{\eta}$ based on the most sensible set of expectations that can be constructed. We will also present robustness checks using other sensible expectations, and we will demonstrate that $\widehat{\frac{\gamma}{\eta}} = 1$ only under implausible beliefs.

Gasoline prices move very closely with crude oil prices: Light Sweet Crude Oil spot prices predict 93 percent 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. A time series of U.S. average retail gasoline price expectations $E_t[g_s]$ is therefore constructed from Light Sweet Crude Oil futures prices from two sources, the Intercontinental Exchange (ICE) and the New York Mercantile Exchange (NYMEX). Table 2 shows the annual average real retail gasoline prices and crude oil futures prices transformed to dollars per gallon of gasoline.

Although oil futures contracts are only traded with high liquidity for settlement dates less than two to three years in the future, Table 2 illustrates that there are some trades observed for settlement dates as far as ten years in the future. The market does not believe that gasoline prices are a martingale: as illustrated in Figure 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. To model expectations for periods beyond the last settlement date observed at each time $t$, our primary specification uses a simple model of mean-reverting expectations, where deviations from a \$1.50/gallon mean decay exponentially. As detailed in the Data Appendix, the mean-reversion parameter is calibrated using all futures data since 1991. The equation fits the data very well: it explains 85% of the variation in the observed futures prices over our 1999-2008 study period.

## 5.2 Reduced Form Data Overview

The dataset we use is perhaps the most wide-ranging data ever assembled in the economics literature on the automobile industry. Before moving to the parameter estimates, we find it useful to give an aggregate, reduced-form overview to build intuition for how the parameters are identified and what the data will show.

Figure 2 illustrates the variation in the instrument $G_{j0(t-a)}$, which will identify price elasticity of demand. For different fuel economy categories, we plot the average value of the instrument for each model year, conditional on time and model year dummies $\tau$ and $\mu$, which in the first stage regression will be the most important covariates, and on the $\psi_{ja}$ fixed effects. Although some identification is generated by the effects of the 1985-1986 gasoline price collapse, the primary source of variation for the first stage is from the gasoline price increase between 2003 and 2008. Over those years, $\Delta G_{j0(t-a)}|(\tau, \mu)$ rises for the lowest MPG (highest GPM) classes, as illustrated by the solid black and dashed blue lines. While unconditional $G_{j0(t-a)}$ also rises for higher MPG vehicles, it rises less than for low MPG vehicles, so their conditional $\Delta G_{j0(t-a)}|(\tau, \mu)$ drops. Within the $ja$ fixed effect groups, this instrument will be negatively correlated with sales of new vehicles.

Figure 3 illustrates the variation in $G_{jat}$ over the study period, conditional on the same dummy variables and fixed effects. As gas prices rose in the latter half of the decade, the conditional $\Delta G_{jat}|(\tau, \mu)$ rose for low fuel economy vehicles, as again illustrated by the solid black and dashed blue lines. Observe that the vertical ordering of the lines had been opposite before 2005: higher-MPG vehicles have higher values of $\Delta G_{jat}|(\tau, \mu)$. This is because the analysis looks within vehicle over time, conditional on time dummies: for all vehicles, the within-vehicle values of $\Delta G_{jat}$ between 2001 and 2005 are lower than after 2005, but they are relatively lower for gas guzzlers during the early years of the study period. Another source of variation not shown by this graph comes from the vehicle's age: those with longer remaining lives see larger changes in $G_{jat}$ when gas price expectations change.

Figures 4 and 5 shift attention to the outcome variables of the first and second stages. Figure 4 displays the sum of sales for vehicles below 19 MPG and above 28 MPG over the past ten years. As gas prices rose between the 2004 and 2007 model years, higher MPG vehicles increased in sales by 450,000 units per year, while sales of the lower-MPG vehicles decreased by 1.4 million units per year. This reinforces the intuition for our instrument: there is a greater market share of three-year-old high MPG vehicles now in 2010 than there were in 2007. This variation in available quantities should cause variation in equilibrium prices.

Figure 5 shows the relative prices of two- to five-year old vehicles in the same low and high MPG categories. As spot gasoline prices rose over the study period, the relative prices of high MPG vehicles similarly increased. Relative vehicle prices appear also to be responsive to higher-frequency gasoline price changes, mirroring changes from 2001-2003 and in 2005-2007.

Figure 5 reinforces the intuition for the "reduced form" version of our identification strategy. Most of our identification comes from the gasoline price increase between 2002 and the end of our study period in 2007-2008. Between the beginning and the end of that period, the incremental cost to fuel the average high vs. low MPG vehicle illustrated in the figure increased by $3281, from $4508 to $7790. Meanwhile, the relative price of the low MPG vehicle decreased by $2384, from $6272 to $3889. In this reduced form example, with a particular subset of the market over a particular time period, relative prices adjusted by 73 percent of the change in gasoline costs.

# 6 Results

This section presents the estimation results for the reduced form model in equation (3) and our nested logit model in equation (9). We detail the sensitivity of our findings to a large array of modeling assumptions and explore various explanations for the results. The primary specification uses the nested logit model with vehicle class as its only nest. To avoid having thinly-traded vehicles drive the estimation results, we weight by the number of transactions within $jat$ used to compute $p_{jat}$. The primary specification includes all passenger cars and light trucks age 0 to 25 years from January 1999 through March 2008. It is important to end the primary specification in that month because macroeconomic changes beginning in the second quarter of 2008 had substantial effects on vehicle markets which would be difficult to model convincingly and could have different effects on different types of vehicles.

**First Stage.** Table 3 shows the first stage regression results for the primary specification. The first stage can be viewed as a reduced form relation between new vehicle quantities and expected gasoline costs. The log of market share ($s_{jat}$) is regressed on expected gasoline costs ($G_{jat}$) and the gas price instrument $G_{j0(t-a)}$, both measured in $1,000s. The negative coefficient on the instrument suggests that fewer new low fuel economy vehicles are sold in equilibrium when gas price expectations are higher. This result is qualitatively consistent with analyses of new vehicle sales in recent analyses by Klier and Linn (2008), Li, Timmins, and von Haefen (2009), and Busse, Knittel, and Zettelmeyer (2009).

**Estimation Results.** Table 4 compares different conceptual approaches to estimating $\frac{\gamma}{\eta}$. Column (1) is our primary specification: the instrumental variables estimation of the nested logit model. The market share coefficients indicate that a one percent increase in market share results in a \$24 drop in the vehicle's price, whereas a one percent increase in the market share of other vehicles in the nest is associated with an \$18 drop in price. Together, these suggest that the correlation parameter $\sigma$ is just over 0.4, so that utilities within a class have a moderate correlation. The coefficient on gas cost is the negative of the estimate of $\frac{\gamma}{\eta}$, meaning that $\widehat{\frac{\gamma}{\eta}} = 0.61$. Thus, our primary specification suggests that changes in market equilibria account for 61 percent of gasoline costs.

Column (2) shows the reduced form from equation (3), in which market shares are assumed to be uncorrelated with $G_{jat}$. Our finding that $\widehat{\frac{\gamma}{\eta}} = 0.52$ is consistent with the intuition that this assumption biases the estimate towards zero due to endogenous market shares.

Column (3) illustrates a representative-consumer logit specification, instead of the nested logit. The point estimate of $\frac{\gamma}{\eta}$ is larger than that of the reduced form, as expected. Because the representative-consumer logit overstates substitutability between different vehicle classes, however, it is unexpected that this specification gives a larger $\widehat{\frac{\gamma}{\eta}}$ than the nested logit model.

Column (4) in Table 4 shows the nested logit model, estimated using ordinary least squares instead of the instrumental variables procedure. The discrete choice version of simultaneity bias - the correlation between market shares and unobserved product quality - often causes estimated price elasticities to have the wrong sign, and that is indeed what we observe here. This underscores the importance of our instrument and IV procedure.

**Alternative Nest Structures.** Table 5 explores alternative nest structures, showing that the basic result is not sensitive to alternate assumptions about substitution patterns. Column (2) allows additional correlation in taste errors among vehicles in the same vehicle class and age category.[19]

[19] Stolyarov's (2002) analysis shows that vehicle trade volumes are highest for vehicle ages near five and ten years. We therefore define the age categories to be 0-4 years, 5-10 years, and greater than 10 years.

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) of Table 5 uses a two level nested logit where the broader nest is vehicle class and the narrower 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, age category, and style. While the coefficients on share variables change as the nest structure is changed, $\widehat{\frac{\gamma}{\eta}}$ is quite stable. While these alternative nest structures does not exhaust all possible forms of substitution patterns, they do suggest that uncaptured substitution patterns in the primary specification are unlikely to cause a bias in $\widehat{\frac{\gamma}{\eta}}$. Appendix 3 presents a series of "reality checks" of the predicted substitution patterns, including implied markups and own- and cross-price elasticities for popular vehicles.

**Discount Rates.** Related literature has framed the question as estimating the "implicit discount rate" that sets $\frac{\gamma}{\eta} = 1$. We chose our "attention weight" formulation because we suspect that it is more behaviorally descriptive than an implicit discount rate, especially since the purchase prices of autos and many other energy-using durables are often amortized over time when they are bought on credit. It is useful, however, to compute this parameter for comparability with other studies and to test how much alternative discount rates affect $\widehat{\frac{\gamma}{\eta}}$.

Columns (2)-(4) in Table 6 show estimates of $\frac{\gamma}{\eta}$ when consumers use a selected annual discount rates rather than 9% in the primary specification. Consistent with intuition, consumers appear more sensitive to gasoline costs when the higher discount rate attenuates the constructed $G_{jat}$ variable. The rates in columns (3) and (4) of Table 6 are chosen to show that $\widehat{\frac{\gamma}{\eta}}$ reaches unity only when the discount rate reaches 27%, while $\widehat{\frac{\gamma}{\eta}}$ is statistically indistinguishable from one when the annual discount rate is as low as 18%. Although some subprime borrowers face interest rates at these levels,[20] we believe that the upper bounds of the average interest rate benchmarks discussed

[20]For example, the mean annual nominal interest rate on very deep subprime loans described in Adams, *et al.* (2009) is 26.2%.

in Section 5.1.2 are appropriately conservative in the primary specification.

**Static Model.** In the literal interpretation of a static discrete choice model such as ours, consumers choose a vehicle in each period, but they expect to hold the vehicle for the rest of its useful life. In reality, the consumer's true problem is dynamic: at every point in time, she will have the opportunity to re-sell the vehicle and purchase a new one. A consumer who resells her vehicle has smaller gasoline expenditures than $G_{jat}$, which is computed based on the expected fuel costs over the life of the vehicle. However, the resale price that the consumer receives incorporates the fuel costs over the remaining vehicle life after resale. Thus, regardless of whether the model is static or dynamic, the purchase price should reflect the full stream of future gasoline costs. We show this more formally in Appendix A.1, which describes the assumptions required to simplify from a dynamic model to the static model.

In Appendix A.1, the crucial assumption required to show that our static model is an unbiased simplification of the dynamic problem is a weak form of stationarity: we must assume that when forming beliefs about future resale prices, consumers believe that changes in gasoline prices are uncorrelated with changes in future market shares. The stationarity assumption allows us to substantially simplify how consumers form expectations of a vehicle's future resale price. In practice, market shares do respond to gas prices, and as discussed earlier in the paper, the effect of this response is to attenuate the effect of gas price changes on relative vehicle price changes. If consumers anticipate this future quantity response, a gas price increase should not decrease willingness to pay for a gas guzzler as much as our model predicts. This simplification is not conservative: it biases our estimator toward zero.

Perhaps the most aggressive test of the importance of this issue is to examine whether consumers appear to fully value gas costs accrued only during the time that they will own the vehicle, regardless of the resale value. Using information from Stolyarov (2002), we calculate that the median vehicle holding period is five years. As shown in column (5) of Table 5, $\widehat{\frac{\gamma}{\eta}} < 1$ even when $G_{jat}$ is computed only over a five-year time horizon.

**Intensive Margin.** We account for the short-run elasticity of vehicle use to gasoline prices as discussed during the formulation of $G_{jat}$ and detailed in Appendix 2. Column (2) of Table 7 presents the results of this alternative specification. The estimated $\widehat{\frac{\gamma}{\eta}}$ does not change significantly, although the standard error of the estimate increases.

**Alternative Gasoline Price Expectations.** While the primary specification arguably includes the most defensible formulation of gasoline price expectations, we now present alternative formulations. In column (3) of Table 7, consumers are assumed to believe that gas prices are a martingale, so that any change in spot gas prices is assumed to be permanent. By overstating (relative to the futures market) how gasoline price changes affect changes in $G_{jat}$ over time, the martingale assumption would bias the model to expect larger changes in relative vehicle prices than the market should actually generate. As illustrated by the regression results, this strengthens the rejection of the null hypothesis.

One can also determine the degree of mean reversion that consumers expect in gasoline price changes that is consistent with the rational model. Column (4) in Table 7 shows that this implicit mean reversion constant is -0.29. This is substantially larger than the -0.057 suggested by the futures market data. Loosely put, auto consumers' beliefs about future gasoline prices would need to have been very strange to explain our results.[21]

**Sticky Information.** A burgeoning literature in macroeconomics, including Reis (2006) and Mankiw and Reis, (2002, 2006), models consumers and firms that face costs in updating information as they choose consumption plans and prices. As a result, consumption and prices do not immediately and fully adjust to news, and they are sensitive to past news. Our model is identified off of variation in gasoline price expectations at a monthly frequency. Although most people presumably drive by a gas station with posted gasoline prices at some point during the choice process, there is a frequency at which it would be unrealistic to expect vehicle market participants to update

[21] While many automobile consumers may not be aware of oil futures prices, information about the oil market's expectations is likely to be transmitted to consumers through the news media, such as in Krueger (2005).

We believe that the assertion that the auto market's expectations of oil prices differs substantially and predictably from the oil market's expectations would be equally remarkable as our interpretation, which is that consumers form expectations on average *as if* informed by futures markets but undervalue fuel costs in choosing between automobiles.

information. In theory, both inattention to future fuel costs and inattention to "high-frequency" fluctuations in gasoline prices could explain our results.[22]

We test for sticky information by including changes in expected gasoline costs over recent periods: $G_{ja(t-s)} - G_{jat}$, with $s$ =1, 4, and 12 months. In the null, where information is not sticky, these changes should not be correlated with current prices. As shown in Table 8, these changes are significantly correlated with $p_{jat}$, which provides empirical support for sticky information. Controlling for these recent changes, however, does not statistically change $\widehat{\frac{\gamma}{\eta}}$, and the estimate remains significantly less than unity.

**Sensitivity to Time Period and "Extraordinary" Events.** In principle, it is possible that $\frac{\gamma}{\eta}$ has changed over time. In practice, our estimation strategy is limited by the fact that sufficient variation in gas prices is needed both during the period when the observed vehicles were produced (for power in the first stage) and during the period of study (for power in the second stage). In Table 9, we break the sample into earlier and later periods, 1999-2005 and 2004-2008.[23] In column (2), $\widehat{\frac{\gamma}{\eta}}$ is larger for the early period than the full sample, but also has a larger standard error, most likely due to lower variation in gas prices during that period. In fact, the estimate is statistically indistinguishable from unity and from the primary specification. In column (3), $\widehat{\frac{\gamma}{\eta}}$ is nearly identical for the later time period as for the full sample.

Ending our primary specification after March 2008 comes at the cost of eliminating potentially-useful variation in gasoline price expectations: retail spot prices rose from \$3.07 in March to \$3.61 in July, and the 8-9 year futures rose by almost as much. Furthermore, one might believe

[22] To formalize this, notice that the utility function could have been written with separate parameters $\widetilde{\gamma}$ and $\gamma$ for attention to the vehicle's average future gas costs over time $\overline{G}_{ja}$ and deviations from that average $G_{jat} - \overline{G}_{ja}$:

$$u_{ijat} = \eta(w - p_{jat}) - \gamma(G_{jat} - \overline{G}_{ja}) - \widetilde{\gamma}\overline{G}_{ja} + \psi_{jat} + \epsilon_{ijat} \tag{13}$$

Given the use of fixed effects, $\widetilde{\gamma}$ is not identified. The ideal test between these two explanations would be to have one permanent change to the market's gasoline price expectations with no corresponding changes to the choice set. Estimating the rate of relative vehicle price adjustment over the ensuing time period would give a sense of information diffusion, while estimating $\frac{\gamma}{\eta}$ with vehicle prices after a very long time would be a convincing measure of attention to future gas costs with full information.

[23] The ends of the periods are chosen in order to avoid weak instruments. They are determined such that the F-statistic of the correlation of the excluded instrument with the endogenous variable is greater than 16.38, the critical value such that the maximal size of a Wald test with $\alpha = 0.05$ is less than 0.1, as suggested by Stock and Yogo (2005). This rule led to a two year overlap in the time periods.

that consumers became especially attentive to gasoline prices as they spiked during that year. This would be consistent with existing work that models "extraordinary" events about which information diffuses instantly (Reis 2006) or that cause consumers to update beliefs between coarse categories (Mullainathan 2002) - e.g. from gas costs being "inconsequential" to gas costs being "high." Column (4) of Table 9 repeats the primary specification including April through December 2008. The point estimate of $\widehat{\frac{\gamma}{\eta}}$ is closer to unity, but still significantly different from that value and not statistically different from the primary specification.[24]

**Changing Consumer Preferences and Vehicle Characteristics.** Changes over time in unobserved consumer preferences, as modeled by $\xi_{jat}$, could bias our estimate of $\frac{\gamma}{\eta}$ if they are correlated with $G_{jat}$. A leading suggestion is that consumers became increasingly "green," or environmentally-oriented, over the study period, resulting in increased preference for high fuel economy vehicles independent of the financial savings. This example causes a bias away from concluding that consumers undervalue gasoline costs: it is an unaccounted effect that should increase the relative price of high fuel economy vehicles over time, while we find that the relative price of high fuel economy vehicles did not increase as much as the model predicts. Columns (2) and (3) in Table 10 show that excluding hybrids or vehicles rated as the most "green"[25] indeed cause the estimated $\widehat{\frac{\gamma}{\eta}}$ to move slightly further away from unity.

Similarly, our estimator could be biased if changes in unobserved vehicle characteristics within a model $j$ over model years are correlated with $G_{jat}$. For example, we have limited statistical evidence (available upon request) that manufacturers changed amenities, such as engine displacement, torque, wheelbase, and stability control, differentially within low-vs. high MPG models over time. We address this partially by defining vehicles as different models $j$ if their engine displacement or fuel economy change by more than ten percent between model years. Perhaps the best suggestive test of the importance of these concerns is to add controls for all observed characteristics, which still have some small residual variance conditional on the fixed effects. Columns (4) and (5) of

[24] In a similar vein, we find in an alternate specification (not shown) that each $1 increase in gas costs has a statistically significant 1.5 cent greater effect on prices in months when retail gas prices changed by at least ten cents per gallon. This provides some evidence that consumers may be more attentive during periods of larger change, but the difference is economically small and does not appear to explain our primary findings.

[25] Defined as the top 60 vehicles in Yahoo's ranking of the 100 greenest vehicles, at http://autos.yahoo.com/green_center-top100/

Table 10 show that these controls have almost no impact on the results, suggesting that most likely any correlation between gas costs and unobserved changes in characteristics is small.

A related concern is that unobserved vehicle characteristics are correlated with the instrument $G_{j0(t-a)}$. While again, we cannot directly test the exclusion restriction, we can test whether within-model observable characteristics affect the first stage relationship. The respective first stages (not shown) of the specifications in columns (4) and (5) of Table 10 similarly indicate that the additional controls have almost no impact on the first stage coefficients.

**Used Vehicle Retail Prices from JD Power Dealership Data.** Because we use fixed effects, as long as changes in $G_{jat}$ are passed through from wholesale to resale prices in levels, the use of wholesale vs. retail data should not affect our estimated $\widehat{\frac{\gamma}{\eta}}$. As an additional specification check, however, we estimate the model using retail transaction prices for used vehicles, which are also included in the JD Power dealership data. Many of these are the same vehicles that went through the Manheim auctions. We use the Manheim wholesale data as the measure of used vehicle prices in our primary specification because these data include substantially more $jat$ observations (over one million vs 500,000), while there are fewer than 1000 observations in JD Power that are not in Manheim.

As shown in Table 11, retail prices predict a smaller $\widehat{\frac{\gamma}{\eta}}$ than do auction prices. Column (2) repeats the primary specification, except limited to the sample of vehicles that are common to both data sets. Column (3) shows the same sample with the JD Power used vehicle prices. Column (4) shows the full JD Power sample. These suggest that retail-wholesale markups for used vehicles are actually slightly negatively correlated with gas costs $G_{jat}$ within vehicle, and thus our primary specification may slightly overstate consumers' $\widehat{\frac{\gamma}{\eta}}$.

**Measurement Error.** If the gas cost variable $G_{jat}$ is measured with error, the estimate of $\widehat{\frac{\gamma}{\eta}}$ may be affected by attenuation bias.[26] The three components of this constructed variable that are most

[26] Measurement error is a separate issue from heterogeneity in gas costs across consumers. While this heterogeneity certainly exists, our results are consistent if we observe without error the gas costs of a "marginal" individual who determines the relative price of a vehicle. Furthermore, as discussed in section 4.2, we intentionally introduced conservative biases in $G_{jat}$. Because these are consistent biases in variable construction, they are distinct from the type of measurement error considered in this subsection.

likely to be measured with error are gasoline prices, miles per gallon, and vehicle miles traveled. We examine each of these separately by instrumenting for gasoline costs $G_{jat}$ with a quantity which is correlated with gas costs, but (we hope) not the econometric error. Each instrument will be constructed similarly to $G_{jat}$ with equation (12), but with one difference.

To address potential measurement error in gasoline prices, our instrument is $G_{jat}$ calculated with lagged gasoline prices and price expectations. Columns (2) and (3) of Table 12 show the results when the instrument is calculated with a one month and one year lag, respectively. To address measurement error in fuel economy, our instrument is $G_{jat}$ computed with the average inverse fuel economy (gallons per mile) across a model and age. Column (4) shows the result with this instrument. To address measurement error in vehicle miles traveled, our instrument is $G_{jat}$ computed with an annual VMT of 12,000 miles for all vehicles. Column (5) shows the result with this instrument. $\widehat{\frac{\gamma}{\eta}}$ is not significantly different from the primary specification in any of these columns.

## 6.1 Magnitude of Mispricing

What are the real-world magnitudes of our parameter estimates? For a set of example vehicles, we now illustrate how much the null hypothesis predicts that vehicle prices should change, compared to how much they do change.

Recognize first that our empirical approach is not informative about the absolute mispricing of each vehicle at a given time. Our coefficient estimates do predict, however, how much the relative prices of vehicles with different fuel economy ratings change in response to a given change in gasoline price expectations. Consider a set of example used vehicles that will hypothetically be driven 12,000 miles per year for the remaining seven years of their lifetimes. The "Predicted Price Change" line in Figure 7 illustrates changes in relative prices caused by a permanent \$1 increase in gasoline prices. This line is computed using the change in $G_{jat}$ for these example used vehicles and the $\widehat{\frac{\gamma}{\eta}}$ from our primary specification. The Honda Civic, which has fuel economy of 30 miles per gallon, is normalized to have zero relative price change. We hold market shares constant, so this can be viewed as a short-run effect on prices.

As shown in the graph, our parameter estimates predict that a Ford F-150, rated at 15 MPG, with this assumed VMT and remaining lifetime would see its relative price drop by \$1010 compared to the Honda Civic. The double line on the graph presents the relative price changes that would be expected if $\frac{\gamma}{\eta} = 1$; the F-150's relative price should drop by \$1660. This \$650 difference suggests that vehicles were substantially mispriced as gasoline prices changed over the study period.

## 7 Welfare Implications

A leading explanation for our empirical result is that consumers misoptimize: they are less attentive to gasoline costs than to purchase prices. For this section, we take as given that this argument is accepted. What, then, are the welfare gains from policies to correct this misoptimization? This section defines the hedonic utility function, develops a new and highly tractable approach to behavioral welfare analysis in a discrete choice setting, and presents the welfare results.

We distinguish between *choice utility*, the utility function that the consumer maximized at the time of choice, and *hedonic utility*, the utility that the consumer actually experienced as a result of the choice. The original utility function in equation (6) was the choice utility function, as it was used to parameterize a demand model estimated off of consumers' observed choices. For rational consumers, choices maximize hedonic utility, and choice utility and hedonic utility are equivalent. If we accept that choices do not reveal hedonic utility, however, we must take an alternative stand on how to construct hedonic utility functions. Our approach is to maintain the assumption that a fully-optimizing agent would have $\gamma = \widehat{\eta}$: consumers' responses to purchase price variation reflect their true marginal utility of money.[27][28] The hedonic utility function is:

[27] We note that our results are also theoretically consistent with consumers that attend correctly to gasoline costs but are *over*-attentive to purchase price relative to their private optimum. Interestingly, if this were the case, correcting the internality would cause consumers to be less price elastic and thus to buy more expensive cars, which on average have lower fuel economy. However, the idea that consumers correctly perceive product price but can misperceive other costs due to sales tax, future gasoline prices, add-ons, and shipping and handling appears to be the most natural interpretation of the results of our paper and related work.

[28] Our approach can be thought of as an application of Bernheim and Rangel (2009) to the case of misoptimization in discrete choice models. In their language, vehicle purchase is a "Generalized Choice Situation" in which consumer $i$ chooses between a set of vehicles with total discounted user costs $p_{jat} + G_{jat}$ and utility flows $\psi_{jat} + \epsilon_{ijat}$. Whether the total cost flows through $p$ or $G$ is an "ancillary condition," meaning that while it may affect choices by agents who misoptimize, it is not material to welfare. We estimate elasticity to total discounted user cost from only the "welfare-relevant domain," which we assume to be only the variation generated by variation in purchase prices. Conversely, we assume that variation in total discounted user cost resulting from variation in $G$ is "suspect," meaning that it should not be used to infer utility functions.

$$u_{ijat} = \eta(w - p_{jat} - G_{jat}) + \psi_{jat} + \epsilon_{ijat} \tag{14}$$

We note that, as in the empirical analysis, the problem has been simplified by assuming zero wealth effects. Consumer $i$'s hedonic utility $u^h_{ijat}$ from choice $jat$ can be written as choice utility $u^c_{ijat}$ minus an "internality" $u^b_{ijat}$:

$$u^h_{ijat} = u^c_{ijat} - u^b_{ijat} \tag{15}$$

In our application, $u^b$ captures the utility value of the portion of future gasoline costs that the consumer did not appropriately value in the discrete choice. This can be thought of as consumption of the numeraire good that the consumer anticipated having at the time of the discrete choice, but does not actually have because of additional expenditures on gasoline. Subtracting $u^h$ from $u^c$, we have:

$$u^b_{ijat} = (\eta - \gamma)G_{jat} \tag{16}$$

The analytical appeal of this approach is that we have written hedonic utility as the sum of two terms that can be integrated over consumers. Summing over the choices made by consumers of market size $M$ and transforming from utility to dollar terms by dividing by $\eta$, we have the expected internality $CS^b$:

$$CS^b = \frac{1}{\eta} \cdot \frac{1}{M} \sum_{i=1}^{M} u^b_i = (1 - \frac{\gamma}{\eta})G_{jat} \cdot s_{jat} \tag{17}$$

Define the variable $\delta_{jat} = -\eta p_{jat} - \gamma G_{jat} + \psi_{jat}$ as the average choice utility for product $ja$ at time $t$. We integrate up over the logit error to get the expectation of "choice consumer surplus" using well known formulas originally from Small and Rosen (1981), modified for the nested logit. Omitting the constant for exposition, we have:

$$CS^c = \frac{1}{\eta} \ln \left[ \sum_{n \in \mathcal{N}} \left[ \sum_{ja \in \mathcal{B}_n} \exp\left( \frac{\delta_{jat}}{1-\sigma} \right) \right]^{1-\sigma} \right] \tag{18}$$

Having defined the analytical approach to welfare analysis of policy changes, consider now a "Behavioral Feebate" policy designed to move consumers purchasing new vehicles to their private optima. The Behavioral Feebate imposes a fee on new vehicle purchases of the fraction of lifetime gas costs that consumers appear to undervalue, $(1 - \widehat{\frac{\gamma}{\eta}}) \cdot G_{jat}$, while rebating some amount $R$. The amount of the Feebate $F_{jat}$ is:

$$F_{jat} = \left\{ \begin{array}{cc} \left(1 - \frac{\gamma}{\eta}\right) \cdot G_{jat} - R, & a = 0 \\ 0, & a > 0 \end{array} \right\} \tag{19}$$

Observe that when this $F_{jat}$ is substituted into new vehicle price in the equation for average choice utility $\delta_{jat}$, it produces a coefficient of $\eta$ on $G_{jat}$:

$$\begin{aligned} \delta_{jat} &= -\eta \left[ p_{jat} + \left(1 - \frac{\gamma}{\eta}\right) \cdot G_{jat} - R \right] - \gamma G_{jat} + \psi_{jat} && (20) \\ &= \eta(w - p_{jat} - G_{jat} + R) + \psi_{jat} && (21) \end{aligned}$$

The choice utility function now equals the hedonic utility function, with the addition of $R$, which modifies the price level. In our Behavioral Feebate counterfactual, we choose to use the $R$ such that the policy leaves unchanged the aggregate market share of new vehicles. Put differently, this assumes that consumers will not perceive a change in new vehicle price level from the policy, only a change in relative prices across vehicles. Because this policy need not be revenue-neutral, to compute the change in choice consumer surplus, any deficit or surplus funds are recycled to all consumers (including those who purchase the outside option) as a lump sum tax or subsidy.

We recognize that under the counterfactual policy, many aspects of vehicle markets would be different. For example, relative prices of used high fuel economy vehicles would increase, and auto manufacturing firms would likely offer more high MPG vehicles and invest more in R&D to improve fuel economy. Simulating these effects is well beyond the scope of this paper. Our simulation assumes that the prices and characteristics of the year 2007 choice set are constant, adds $F_{jat}$ to vehicle prices, and resimulates market shares.

The fitted average utility from vehicle ownership $\widehat{\psi}_{jat}$ is backed out from the observed baseline market shares $s_{jat}$ and the estimated demand parameters $\widehat{\eta}$, $\widehat{\gamma}$, and $\widehat{\sigma}$, using equation (7). As in the estimation, the outside option is its own nest, with utility normalized to zero. Intuitively, welfare gains flow through the fact that the feebate causes consumers to spend less money on gasoline and more on some combination of higher fuel economy vehicles and the numeraire good.

## 7.1 Welfare Analysis Results

Table 13 shows the simulation results. All new vehicles with fuel economy below the "pivot" of 19.0 MPG see an increase in sales, while those with fuel economy below 19.0 MPG see a decrease. The average fuel economy of the new vehicle fleet increases by 2.36 miles per gallon.

Choice consumer surplus $CS^h$ drops by a net present value of \$17 per potential vehicle consumer per year of the Behavioral Feebate policy, as the policy moves consumers away from their perceived optimum. The average internality $CS^b$, however, is \$32 lower. Hedonic consumer surplus therefore increases by \$15, as consumers buy higher fuel economy vehicles, total gasoline costs drop, and consumers spend more on the numeraire good.

The policy reduces gasoline use by each year's new vehicle consumers by 37.5 gallons over the lives of their vehicles. Assuming, for the sake of argument, that the marginal damages of carbon dioxide emissions are \$30 per metric ton, this translates into a reduction in climate damages of \$7.60 in present discounted value.

Comparing these figures shows that the welfare gains from reducing negative externalities are smaller than the welfare gains from reducing the "internality" by inducing consumers to make the *privately*-optimal choice. Intuitively, if we believe that the externality is much less per gallon than the current price, misoptimization over future fuel costs reduces the consumer's own private utility more than it hurts external social welfare. An important takeaway from this analysis, then, is that behavioral misoptimization can be a more powerful justification for energy efficiency and fuel economy standards than internalizing environmental externalities. In a theoretical sense, of course, this was clear *a priori*: if consumers did not misvalue future gasoline costs, there is no economic argument for adding a fuel economy standard or feebate to the optimal Pigouvian energy tax. While environmental externalities have often been the center of the discussion of energy efficiency standards, feebates, and fuel economy standards, this simple analysis suggests that understanding consumer choice is much more important from a welfare perspective.

## 8 Conclusion

For more than 30 years, economists have attempted to estimate "implicit discount rates": how much weight consumers place on future energy costs (or other future costs) when they buy energy-consuming durable goods. Building on this literature, this paper tests whether vehicle prices and market shares respond to changes in gasoline price expectations in a way that is consistent with consumers who value equally \$1 in purchase price and \$1 in present discounted fuel costs. We use a discrete choice demand model that addresses several economic and econometric challenges, and we introduce a new instrument for vehicle prices into the empirical literature on automobile demand. The results show that vehicle market equilibria underadjust to changes in expected future gas costs: prices and market shares move as if consumers are willing to pay only \$0.61 up front to reduce discounted gasoline costs by \$1.

The estimated responsiveness of vehicle prices and shares to expected gas costs depends on assumptions about expectations of future gas prices, vehicle lifetime, consumers' discount rates, substitution patterns, and other parameters. We show, however, that market equilibria move as predicted by the null hypothesis only under implausible sets of assumptions. We explore a variety of explanations for our results, including risk aversion, measurement error, sticky information, credit constraints, changes in consumer preferences over time, and other factors. A plausible explanation for the empirical results appears to be that gasoline costs are a "shrouded attribute," and consumers attend to them less than upfront prices at the time of purchase.

This explanation implies that two types of policies could theoretically increase welfare. The first are information provision or marketing programs that nudge consumers toward attending to fuel economy at the time of choice. While these policies are relatively palatable to economists, it is not obvious that their effects are large. For example, auto dealerships are already required to post fuel economy labels on the windows of all new vehicles on their sales lots.

A second type of policies move from "soft paternalism" to paternalism. Feebates, gas guzzler taxes, and Corporate Average Fuel Economy Standards tax or limit the sale of high fuel economy vehicles. In a simple welfare calculation such as ours, these substantially increase welfare, by increasing purchases of high fuel economy vehicles with future financial benefits that consumers undervalued relative to their private optimum.

While our results may provide some economic justification for these policies, this should be viewed very tentatively. As an example of why, notice that our empirical analysis and welfare calculation assumed for simplicity that the misvaluation of future fuel costs is homogenous in the population. In reality, this parameter presumably varies across individuals and over time: some consumers may overvalue fuel economy in their decisions, while others undervalue it, and this may change depending on media coverage of gas prices, automakers' sales practices and advertising, or other factors. "Behavioral Feebates" and fuel economy standards are blunt instruments in that they can only generate optimal vehicle choices for one level of misvaluation. As a result, any given policy of this type will be too large for some consumers and too small for others relative to the optimum. Our future work, using randomized trials and natural experiments as well as applied theory, aims to confirm these empirical results, shed additional light on potential explanations, and formalize the implications for business strategy and policy.

# References


[1] Adams, William, Liran Einav, and Jonathan Levin (2009). "Liquidity Constraints and Imperfect Information in Subprime Lending." *American Economic Review*, Vol. 99, No. 1, pages 49–84.

[2] Austin, David (2008). "Effects of Gasoline Prices on Driving Behavior and Vehicle Markets." U.S. Congressional Budget Office Working Paper (January).

[3] Barber, Brad, Terrance Odean, and Lu Zheng (2005). "Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows." *Journal of Business*, Vol. 78, No. 6 (November), pages 2095-2120.

[4] Bento, Antonio, Lawrence Goulder, Mark Jacobsen, and Roger von Haefen (2009). "Distributional and Efficiency Impacts of Increased U.S. Gasoline Taxes." *American Economic Review*, American Economic Review, Vol. 99, No. 3 (June), pages 667-699.

[5] Bernheim, Douglas, and Antonio Rangel (2009). "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *Quarterly Journal of Economics*, Vol. 124, No. 1 (February), pages 51-104.

[6] Berry, Steven (1994). "Estimating Discrete Choice Models of Product Differentiation." *RAND Journal of Economics*, Vol. 23, No. 2 (Summer), pages 242-262.

[7] Berry, Steven, James Levinsohn, and Ariel Pakes (1995). "Automobile Prices in Market Equilibrium." *Econometrica*, Vol. 63, No. 4 (July), pages 841-890.

[8] Berry, Steven, James Levinsohn, and Ariel Pakes (2004). "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy*, Vol. 112, No. 1 (February), pages 68-105.

[9] Busse, Meghan, Christopher Knittel, and Florian Zettelmeyer (2009). "Pain at the Pump: How Gasoline Prices Affect Automobile Purchasing." Working Paper, Northwestern University (March).

[10] Chetty, Raj, Adam Looney, and Kory Kroft (2009). "Salience and Taxation: Theory and Evidence." *American Economic Review*, Vol. 99, No. 4 (September), pages 1145-1177.

[11] Davis, Lucas, and Lutz Kilian (2009). "Estimating the Effect of a Gasoline Tax on Carbon Emissions." National Bureau of Economic Research Working Paper 14685 (January).

[12] DeCanio, Stephen (1993). "Barriers within Firms to Energy-Efficient Investments." *Energy Policy*, Vol. 21 (September), pages 906-914.

[13] Dreyfus, Mark, and Kip Viscusi (1995). "Rates of Time Preference and Consumer Valuations of Automobile Safety and Fuel Efficiency." *Journal of Law and Economics*, Vol. 38, No. 1, pages 79-98.

[14] Ellison, Glenn (2005). "A Model of Add-on Pricing." *Quarterly Journal of Economics*, Vol. 120, No. 2 (May), pages 585-638.

[15] Espey, Molly, and Santosh Nair (2005). "Automobile Fuel Economy: What is it Worth?" *Contemporary Economic Policy*, Vol. 23, No. 3 (July), 317-323.

[16] Finkelstein, Amy (2009). "E-ZTAX: Tax Salience and Tax Rates." *Quarterly Journal of Economics*, Vol. 124, No. 3 (August), pages 969-1010.

[17] Fischer, Carolyn, Winston Harrington, and Ian Parry (2007). "Do Market Failures Justify Tightening Corporate Average Fuel Economy (CAFE) Standards?" *The Energy Journal*, Vol. 28, No. 4, pages 1-30.

[18] Gabaix, Xavier, and David Laibson (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, Vol. 121, No. 2, pages 505-540.

[19] Goldberg, Pinelopi (1995). "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry." *Econometrica*, Vol. 63, No. 4 (July), pages 891-951.

[20] Goldberg, Pinelopi (1998). "The Effects of the Corporate Average Fuel Economy Standards in the US." Journal of Industrial Economics, Vol. 46, pages 1-33.

[21] Greene, David (2007). "Why CAFE Worked." *Energy Policy*, Vol. 26, No. 8, pages 595-613.

[22] Greene, David, Janet Hopson, R. Goeltz and Jia Li, (2007). "Analysis of In-Use Fuel Economy Shortfall Based on Voluntarily Reported Mile-per-Gallon Estimates," *Transportation Research Record*, No. 1983, pages 99-105.

[23] Greene, David, Philip Patterson, Margaret Singh, and Jia Li (2005). "Feebates, Rebates, and Gas-Guzzler Taxes: A Study of Incentives for Increased Fuel Economy." *Energy Policy*, Vol. 33, No. 6 (April), pages 757-775.

[24] Hausman, Jerry (1979). "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables." *Bell Journal of Economics*, Vol. 10, No. 1, pages 33-54.

[25] Hausman, Jerry, and Paul Joskow (1982). "Evaluating the Costs and Benefits of Appliance Efficiency Standards." *American Economic Review*, Vol. 72, No. 2 (May), pages 220-225.

[26] Hossain, Tanjim, and John Morgan (2006). "...Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay." *Advances in Economic Analysis and Policy*, Vol. 6.

[27] Jacobsen, Mark (2008). "Evaluating U.S. Fuel Economy Standards In a Model with Producer and Household Heterogeneity." Working Paper, University of California at San Diego.

[28] Jaffe, Adam, and Robert Stavins (1994). "The Energy Paradox and the Diffusion of Conservation Technology." *Resource and Energy Economics*, Vol. 16, pages 91-122.

[29] Kahn, James (1986). "Gasoline Prices and the Used Automobile Market: A Rational Expectations Asset Price Approach." *Quarterly Journal of Economics*, Vol. 101, No. 2 (May), pages 323-340.

[30] Kilian, Lutz and Eric Sims (2006). "The Effects of Real Gasoline Prices on Automobile Demand: A Structural Analysis Using Micro Data," Working Paper, University of Michigan (April).

[31] Klier, Thomas, and Joshua Linn (2008). "The Price of Gasoline and the Demand for Fuel Efficiency: Evidence from Monthly New Vehicles Sales Data." Working Paper, University of Illinois at Chicago (September).

[32] Krueger, Alan (2008). "Why the Tepid Response to Higher Gasoline Prices?" *The New York Times*, October 13, 2005.

[33] Langer, Ashley, and Nathan Miller (2009). "Automakers' Short-Run Responses to Changing Gasoline Prices and the Implications for Energy Policy." Working Paper, University of California at Berkeley (September).

[34] Li, Shanjun, Christopher Timmins, and Roger von Haefen (2009). "Do Gasoline Prices Affect Fleet Fuel Economy?" *American Economic Journal: Economic Policy*, Vol. 1, No. 2 (August), pages 113-137.

[35] Manheim (2009). "Used Car Market Report." 14th Edition. Atlanta, GA: Manheim Consulting.

[36] Mankiw, Gregory, and Ricardo Reis (2002). "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve." *Quarterly Journal of Economics*, Vol. 117, No. 4 (November), pages 1295-1328.

[37] Mankiw, Gregory, and Ricardo Reis (2006). "Pervasive Stickiness." *American Economic Review*, Vol. 96, No. 2 (May), pages 164–169.

[38] Metcalf, Gilbert, and Kevin Hassett (1999). "Measuring the Energy Savings from Home Improvement Investments: Evidence from Monthly Billing Data." *Review of Economics and Statistics*, Vol. 81, No. 3, pages 516-528.

[39] Mullainathan, Sendhil (2002). "Thinking Through Categories." Working Paper, MIT (April).

[40] Murtishaw, Scott, and Jayant Sathaye (2006). "Quantifying the Eξect of the Principal-Agent Problem on US Residential Energy Use." Lawrence Berkeley National Laboratory Working Paper 59773 Rev (August).

[41] National Highway Traffic Safety Administration (NHTSA) (2009). "Preliminary Regulatory Impact Analysis: Corporate Average Fuel Economy for MY 2012-MY 2016 Passenger Cars and Light Trucks." Office of Regulatory Analysis and Evaluation, National Center for Statistics and Analysis (August).

[42] Ohta, Makoto, and Zvi Griliches (1986). "Automobile Prices and Quality: Did the Gasoline Price Increases Change Consumer Tastes in the U.S.?" *Journal of Business and Economic Statistics*, Vol. 4, No. 2 (April), pages 187-198.

[43] Parry, Ian, Winston Harrington, and Margaret Walls (2007). "Automobile Externalities and Policies." *Journal of Economic Literature*, Vol. 45, pages 374-400.

[44] Petrin, Amil (2002). "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy*, Vol. 110, No. 4 (August), pages 705-729.

[45] Reis, Ricardo (2006). "Inattentive Consumers." *Journal of Monetary Economics*, Vol. 53, No. 8 (November), pages 1761-1800.

[46] Rust, John (1985). "Stationary Equilibrium in a Market for Durable Assets." *Econometrica*, Vol. 53, No. 4, pages 783-805.

[47] Sallee, James, Sarah West, and Wei Fan (2009). "The Effect of Gasoline Prices on the Demand for Fuel Economy in Used Vehicles: Empirical Evidence and Policy Implications." Working Paper, Macalester College (December).

[48] Sawhill, James (2008). "Are Capital and Operating Costs Weighted Equally in Durable Goods Purchases? Evidence from the US Automobile Market." Working Paper, University of California at Berkeley (April).

[49] Small, Kenneth, and Kurt Van Dender (2007). "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect." *The Energy Journal*, Vol. 28, No. 1, pages 25-51.

[50] Small, Kenneth, and Harvey Rosen (1981). "Applied Welfare Economics with Discrete Choice Models." Econometrica, Vol. 49, No. 1 (January), pages 105-130.

[51] Stock, James, and Motohiro Yogo (2005). "Testing for Weak Instruments in Linear IV Regression." In Stock, James, and Donald Andrews, Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg. Cambridge, England: Cambridge University Press.

[52] Train, Kenneth (1985). "Discount Rates in Consumers' Energy-Related Decisions: A Review of the Literature." *Energy*, pages 1243-1253 (December).

[53] U.S. Bureau of Labor Statistics (2007). "Consumer Expenditure Survey." ftp://ftp.bls.gov/pub/special.requests/ce/standard/2007/quintile.txt

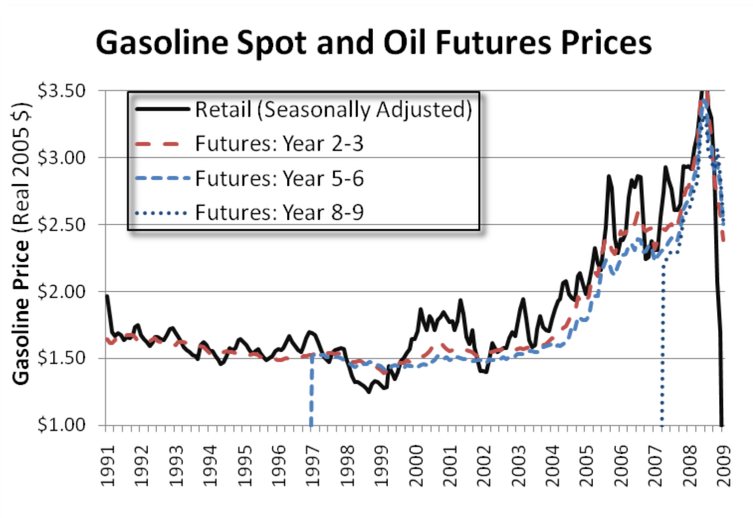
[54] U.S. Department of Transportation (2001). "National Household Travel Survey." http://nhts.ornl.gov/download.shtml#2001.

[55] U.S. Environmental Protection Agency (2008). "Light-Duty Automotive Technology and Fuel Economy Trends: 1975 Through 2008. Appendix A: Database Details and Calculation Methods." http://www.epa.gov/otaq/cert/mpg/fetrends/420r08015appx-a.pdf.

[56] Verboven, Frank (1999). "The Markets for Gasoline and Diesel Cars in Europe." Centre for Economic Policy Research Discussion Paper 2069 (February).

# Figures

## Figure 1: Gasoline Price Expectations




## Figure 2: Identifying Variation in the Instrument

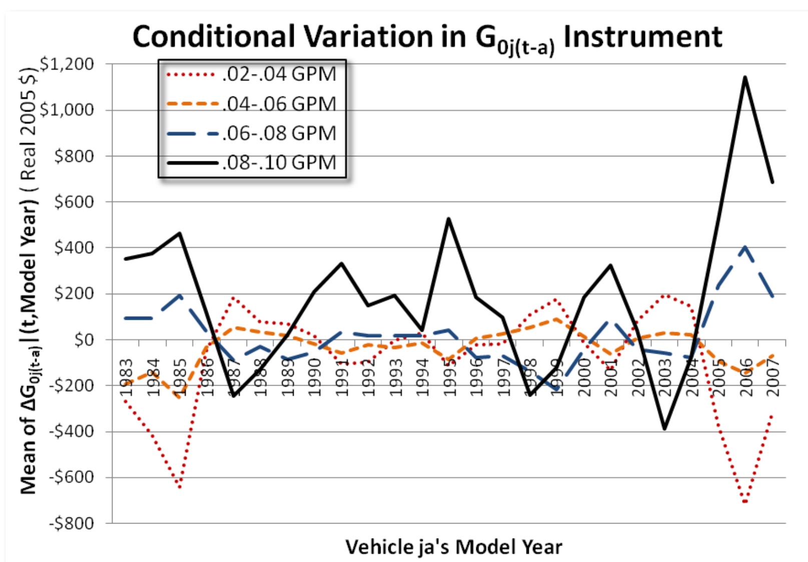


**Figure 3: Identifying Variation in Gasoline Cost**

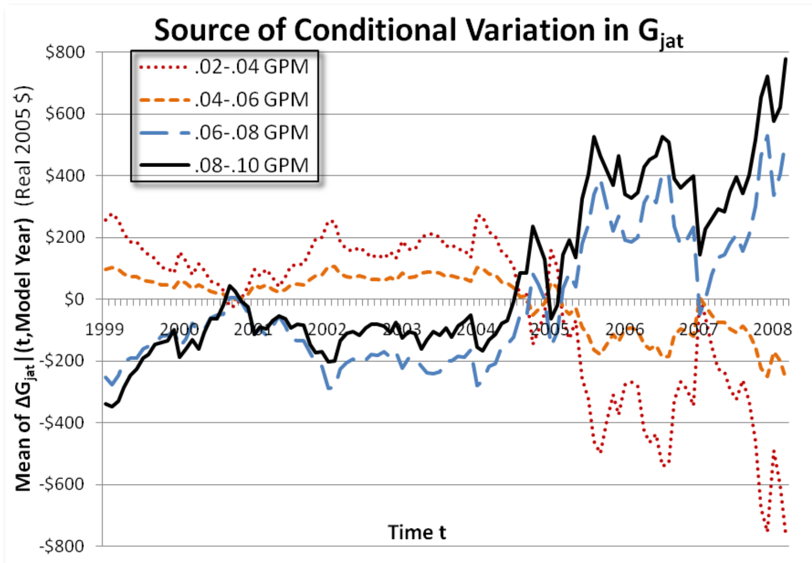



**Figure 4: New Vehicle Sales by MPG Rating**

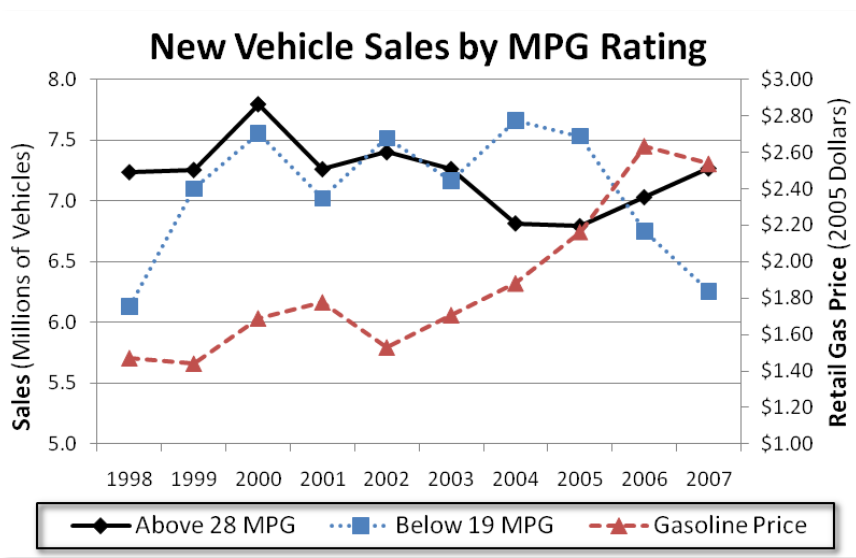


**Figure 5: Relative Prices Low vs. High MPG Vehicles**

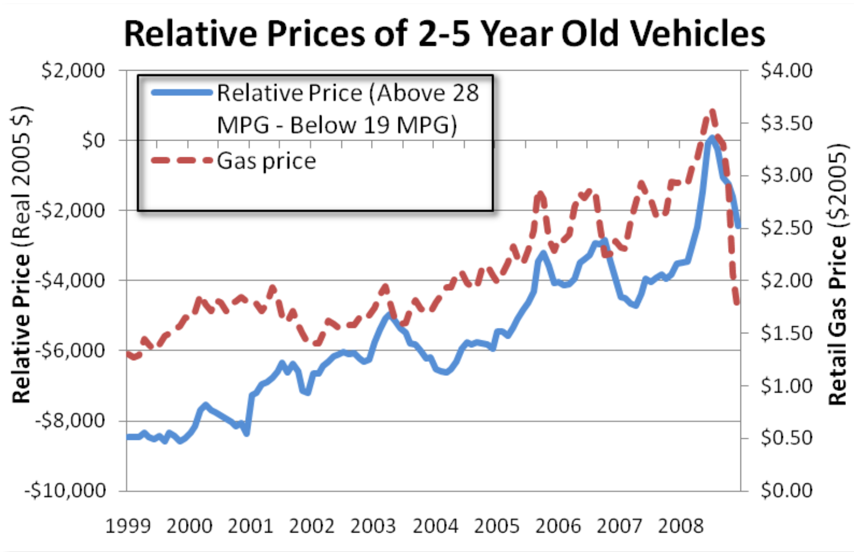
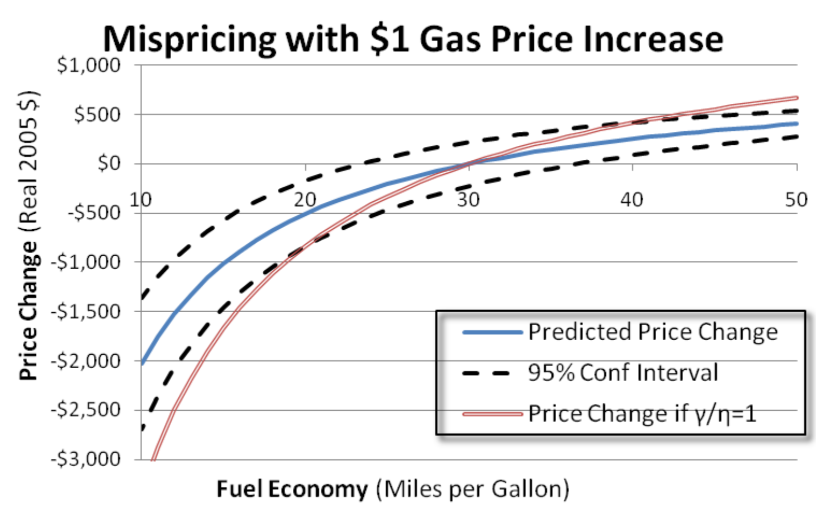



**Figure 6: Mispricing with a $1 Gas Price Increase**



# Tables

Table 1: **Summary Statistics**

| | Full sample | 2007 new models |
|---|---|---|
| Year | 2003.7 | 2007.0 |
| | ( 2.8 ) | ( 0.0 ) |
| Model Year | 1996.4 | 2007.0 |
| | ( 5.6 ) | ( 0.0 ) |
| Age (years) | 7.3 | 0.0 |
| | ( 5.2 ) | ( 0.0 ) |
| Price | 7,863 | 24,872 |
| | ( 8,863 ) | ( 10,383 ) |
| Miles per gallon | 19.1 | 20.8 |
| | ( 4.2 ) | ( 5.1 ) |
| Expected lifetime gas costs | 6,727 | 12,283 |
| (2005 $) | ( 3,794 ) | ( 3,473 ) |
| Horsepower | 236.4 | 260.9 |
| | ( 104.7 ) | ( 77.8 ) |
| Weight (pounds) | 4,296 | 4,617 |
| | ( 1,441 ) | ( 1,035 ) |
| Wheelbase (inches) | 121.8 | 125.8 |
| | ( 19.7 ) | ( 19.8 ) |
| Fraction cars | 0.61 | 0.53 |
| Observations | 1,143,610 | 9,127 |

Notes: Means are quantity-weighted. Standard deviations in parenthesis. The full sample includes monthly observations Jan 1999 - Dec 2008 of all passenger cars and light trucks age 0-25. Column (2) includes 2007 model year vehicles observed in 2007. See text for calculation of expected gas costs.

Table 2: **Gasoline Prices and Expectations**

| Year | **Spot** | **Future Year** 0-1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 | 6-7 | 7-8 | 8-9 | 9-10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1998 | 1.34 | 1.44 | 1.46 | 1.47 | 1.47 | 1.46 | 1.46 | 1.46 | 1.45 | - | - |
| 1999 | 1.43 | 1.50 | 1.46 | 1.45 | 1.44 | 1.44 | 1.43 | 1.43 | 1.39 | - | - |
| 2000 | 1.77 | 1.73 | 1.61 | 1.56 | 1.51 | 1.49 | 1.47 | 1.46 | - | - | - |
| 2001 | 1.69 | 1.65 | 1.59 | 1.55 | 1.52 | 1.51 | 1.49 | 1.48 | 1.47 | - | - |
| 2002 | 1.56 | 1.63 | 1.58 | 1.55 | 1.53 | 1.51 | 1.50 | 1.48 | 1.50 | - | - |
| 2003 | 1.74 | 1.71 | 1.62 | 1.59 | 1.58 | 1.57 | 1.56 | 1.55 | 1.59 | - | - |
| 2004 | 1.99 | 1.95 | 1.84 | 1.78 | 1.74 | 1.70 | 1.68 | 1.66 | 1.75 | - | - |
| 2005 | 2.34 | 2.33 | 2.28 | 2.22 | 2.15 | 2.10 | 2.06 | 2.03 | 2.02 | - | - |
| 2006 | 2.55 | 2.55 | 2.55 | 2.49 | 2.42 | 2.36 | 2.32 | 2.27 | - | - | - |
| 2007 | 2.68 | 2.59 | 2.55 | 2.50 | 2.46 | 2.42 | 2.39 | 2.40 | 2.37 | 2.35 | 2.56 |
| 2008 | 3.00 | 3.12 | 3.10 | 3.08 | 3.06 | 3.04 | 3.02 | 2.99 | 2.97 | 2.95 | 2.67 |

Note: All prices are in dollars per gallon and are inflation adjusted to 2005 dollars. Futures prices are transformed via regression from oil prices to gasoline prices and deflated to 2005 dollars using inflation expectations implied by Treasury Inflation-Protected Security prices.

Table 3: **Nested Logit First Stage**

| | (1) |
|---|---|
| Instrumented variable: | $\ln s_{jat}$ |
| $G_{jat}$ | -0.02 |
| | ( 0.02 ) |
| G at age 0 | -0.13 |
| | ( 0.02 ) |
| Observations | 1,053,058 |
| $ja$ groups | 37,794 |
| R-squared | 0.06 |
| F (excl instruments) | 29.1 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. The instrument 'G at age 0' is the expected gas cost in the year the vehicle was new. All regressors are measured in $1,000s. Standard errors are robust and clustered by $ja$ (model * age).

Table 4: **Comparison of Reduced Form and Nested Logit**

Dependent variable: Vehicle price

| Specification | (1) **Primary** | (2) Reduced form | (3) Logit | (4) Nested Logit OLS |
|---|---|---|---|---|
| $G_{jat}$ | -0.61 | -0.52 | -0.78 | -0.38 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.05 ) | ( 0.09 ) | ( 0.04 ) |
| ln(market share) | -2372 | | -2950 | 179 |
| $[-(1-\sigma)/\eta]$ | ( 723 ) | | ( 734 ) | ( 38 ) |
| ln(nest share) | -1807 | | | -1981 |
| [-$\sigma/\eta$] | ( 655 ) | | | ( 358 ) |
| | | | | |
| Observations | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 |
| *ja* groups | 37,794 | 37,794 | 37,794 | 37,794 |
| F (excl inst) | 29.1 | | 33.8 | |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. 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 is the share of all vehicles in the same class. Standard errors are robust and clustered by *ja* (model * age).

Table 5: **Alternative Nest Structures**

Dependent variable: Vehicle price

| Specification | (1) **Primary** | (2) Class/Age | (3) Age/Class | (4) Class/Style | (5) 3 nests |
|---|---|---|---|---|---|
| $G_{jat}$ | -0.61 | -0.60 | -0.72 | -0.60 | -0.59 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.07 ) | ( 0.08 ) | ( 0.06 ) | ( 0.07 ) |
| ln(share) | -2372 | -2455 | -2766 | -2398 | -2581 |
| $[-(1-\sigma_1)/\eta]$ | ( 723 ) | ( 735 ) | ( 824 ) | ( 743 ) | ( 789 ) |
| ln(nest 1 share) | -1807 | -4598 | -2349 | -2412 | -4365 |
| $[-(\sigma_1-\sigma_2)/\eta]$ | ( 655 ) | ( 1295 ) | ( 1612 ) | ( 831 ) | ( 1324 ) |
| ln(nest 2 share) | | 2304 | -731 | 650 | 1089 |
| $[-(\sigma_2-\sigma_3)/\eta]$ | | ( 805 ) | ( 514 ) | ( 651 ) | ( 878 ) |
| ln(nest 3 share) | | | | | 1071 |
| $[-\sigma_3/\eta]$ | | | | | ( 479 ) |
| | | | | | |
| Observations | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 |
| *ja* groups | 37,794 | 37,794 | 37,794 | 37,794 | 37,794 |
| F (excl inst) | 29.1 | 28.7 | 27.1 | 28.1 | 25.9 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. Column (1) is the primary specification from Table 3, column (1). 'Nest 1 share' is the share of all vehicles in the same class. For models with two nests, 'nest 1 share' denotes the share of all vehicles in the narrowest nest (e.g. the number of vehicles in the same class and age category for column (2)), and 'nest 2 share' denotes the share of all vehicles within the broadest nest. Nest shares are similarly defined for the model with three nests. Column (2) uses two nests, vehicle class, and age buckets (0-4 years, 5-10 years, 11+ years). Column (3) reverses the order of the nests. Column (4) uses as a second nest the "style" of a vehicle (indicators for whether vehicle is a luxury make and whether the firm is based in Europe, North America, or Asia). Column (5) includes all three nests. Standard errors are robust and clustered by *ja* (model * age).

Table 6: **Alternative Discount Rates and Time Horizon**

Dependent variable: Vehicle price

| Specification | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | **Primary** | $r = 5\%$ | $r = 18\%$ | $r = 27\%$ | 5 yr horizon |
| $G_{jat}$ | -0.61 | -0.52 | -0.81 | -1.02 | -0.79 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.06 ) | ( 0.09 ) | ( 0.11 ) | ( 0.09 ) |
| ln(market share) | -2372 | -2442 | -2248 | -2168 | -2029 |
| $[-(1-\sigma)/\eta]$ | ( 723 ) | ( 720 ) | ( 726 ) | ( 728 ) | ( 726 ) |
| ln(nest share) | -1807 | -1804 | -1816 | -1825 | -1842 |
| [-$\sigma/\eta$] | ( 655 ) | ( 663 ) | ( 640 ) | ( 630 ) | ( 612 ) |
| | | | | | |
| Observations | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 |
| *ja* groups | 37,794 | 37,794 | 37,794 | 37,794 | 37,794 |
| F (excl inst) | 29.1 | 30.1 | 27.6 | 26.7 | 25.4 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. Column (1) is the primary specification from Table 3, column (1). Nest share is the share of all vehicles in the same class. A 9% annual discount rate is assumed in the calcuation of gas costs. Columns (2)-(4) use a 5%, 18%, and 27% discount rate in the calcuation of gas costs, respectively. Column (5) uses a 9% discount rate but only accounts for the next 5 years of gas costs. Standard errors are robust and clustered by *ja* (model * age).

Table 7: **Alternate Gas Price Expectations and Vehicle Usage**

Dependent variable: Vehicle price

| Specification | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | **Primary** | Intensive mrgn | Martingale | Mean reversion |
| $G_{jat}$ | -0.61 | -0.35 | -0.44 | -1.00 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.14 ) | ( 0.04 ) | ( 0.15 ) |
| ln(market share) | -2372 | -1749 | -1538 | -4034 |
| $[-(1-\sigma)/\eta]$ | ( 723 ) | ( 785 ) | ( 643 ) | ( 1037 ) |
| ln(nest share) | -1807 | -1742 | -1908 | -2149 |
| [-$\sigma/\eta$] | ( 655 ) | ( 579 ) | ( 544 ) | ( 887 ) |
| | | | | |
| Observations | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 |
| *ja* groups | 37,794 | 37,794 | 37,794 | 37,794 |
| F (excl inst) | 29.1 | 18.5 | 29.1 | 26.7 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. Column (1) is the primary specification from Table 3, column (1). Nest share is the share of all vehicles in the same class. Column (2) assumes that vehicle usage changes with gas prices; see Appendix 2 for a complete description. Column (3) assumes martingale expectations of gas prices. Column (4) assumes mean reverting gas prices with constant -0.29; see Appendix 2 for a complete description. Standard errors are robust and clustered by *ja* (model * age).

Table 8: **Sticky Prices**

Dependent variable: Vehicle price

| Specification | (1) **Primary** | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $G_{jat}$ | -0.61 | -0.66 | -0.68 | -0.64 | -0.63 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.07 ) | ( 0.08 ) | ( 0.09 ) | ( 0.09 ) |
| 1 month lag | | -1.13 | | | -1.07 |
| | | ( 0.08 ) | | | ( 0.08 ) |
| 4 month lag | | | -0.40 | | -0.28 |
| | | | ( 0.08 ) | | ( 0.07 ) |
| 12 month lag | | | | -0.08 | 0.18 |
| | | | | ( 0.08 ) | ( 0.09 ) |
| Observations | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 |
| $ja$ groups | 37,794 | 37,794 | 37,794 | 37,794 | 37,794 |
| F (excl inst) | 29.1 | 28.6 | 26.9 | 24.1 | 23.9 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. Column (1) is the primary specification from Table 3, column (1). Nest share variables are not shown in this table. Columns (2)-(5) add controls for $G_{ja(t-s)} - G_{jat}$, where $s$ is 1, 4, or 12 months. Standard errors are robust and clustered by $ja$ (model * age).

Table 9: **Alternate Time Periods**

Dependent variable: Vehicle price

| Specification | (1) **Primary** | (2) Jan 99 - Dec 05 | (3) Jan 04 - Mar 08 | (4) Jan 99 - Dec 08 |
|---|---|---|---|---|
| $G_{jat}$ | -0.61 | -0.90 | -0.59 | -0.69 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.14 ) | ( 0.05 ) | ( 0.06 ) |
| ln(market share) | -2372 | -4301 | -3447 | -2415 |
| $[-(1-\sigma)/\eta]$ | ( 723 ) | ( 1296 ) | ( 850 ) | ( 690 ) |
| ln(nest share) | -1807 | -602 | -2012 | -1944 |
| [-$\sigma/\eta$] | ( 655 ) | ( 957 ) | ( 859 ) | ( 655 ) |
| Observations | 1,053,058 | 766,713 | 524,093 | 1,143,593 |
| $ja$ groups | 37,794 | 27,825 | 25,976 | 38,534 |
| F (excl inst) | 29.1 | 17.6 | 24.4 | 31.5 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. Column (1) is the primary specification from Table 3, column (1). Nest share is the share of all vehicles in the same class. Columns (2)-(4) limit the sample to the time periods shown. Standard errors are robust and clustered by $ja$ (model * age).

Table 10: **Changing Characteristics and Preferences**

Dependent variable: Vehicle price

| Specification | (1) **Primary** | (2) Exclude hybrids | (3) Exclude 'green' | (4) Char. sample | (5) Char. control |
|---|---|---|---|---|---|
| $G_{jat}$ | -0.61 | -0.61 | -0.49 | -0.57 | -0.57 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.07 ) | ( 0.07 ) | ( 0.09 ) | ( 0.09 ) |
| ln(market share) | -2372 | -2385 | -2617 | -1695 | -1706 |
| $[-(1-\sigma)/\eta]$ | ( 723 ) | ( 727 ) | ( 716 ) | ( 1019 ) | ( 1017 ) |
| ln(nest share) | -1807 | -1807 | -1652 | -132 | -298 |
| $[-\sigma/\eta]$ | ( 655 ) | ( 657 ) | ( 701 ) | ( 1099 ) | ( 1119 ) |
| Observations | 1,053,058 | 1,052,577 | 1,013,868 | 171,873 | 171,873 |
| $ja$ groups | 37,794 | 37,766 | 36,493 | 8,841 | 8,841 |
| F (excl inst) | 29.1 | 29.0 | 29.6 | 14.5 | 14.1 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. Column (1) is the primary specification from Table 3, column (1). Nest share is the share of all vehicles in the same class. Column (2) excludes hybrid vehicles. Column (3) excludes 'green' vehicles. Columns (4) and (5) are limited to a subsample for which additional vehicle characteristics are available, (wheelbase, engine displacement, horsepower, torque, traction control, ABS brakes, and stability control) and column (5) controls for those characteristics. Standard errors are robust and clustered by $ja$ (model * age).

Table 11: **Retail and Wholesale Prices**

Dependent variable: Vehicle price

| Specification | (1) **Primary** | (2) Common sample | (3) Retail prices | (4) Retail only |
|---|---|---|---|---|
| $G_{jat}$ | -0.61 | -0.75 | -0.60 | -0.57 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.09 ) | ( 0.07 ) | ( 0.07 ) |
| ln(market share) | -2372 | -2921 | -2273 | -2174 |
| $[-(1-\sigma)/\eta]$ | ( 723 ) | ( 915 ) | ( 743 ) | ( 719 ) |
| ln(nest share) | -1807 | -1415 | -1775 | -2100 |
| $[-\sigma/\eta]$ | ( 655 ) | ( 742 ) | ( 613 ) | ( 631 ) |
| Observations | 1,053,058 | 475,336 | 475,336 | 476,080 |
| $ja$ groups | 37,794 | 19,039 | 19,039 | 19,095 |
| F (excl inst) | 29.1 | 26.4 | 29.6 | 30.3 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. Column (1) is the primary specification from Table 3, column (1). Nest share is the share of all vehicles in the same class. Column (2) is limited to the sample with both wholesale and retail used vehicle prices available. Column (3) uses the retail price data in the common sample. Column (4) uses the retail price data with the full sample. Standard errors are robust and clustered by $ja$ (model * age).

Table 12: **Measurement Error**

Dependent variable: Vehicle price

| Specification | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | **Primary** | Lag G 1 mo | Lag G 1 yr | G average GPM | G fixed VMT |
| $G_{jat}$ | -0.61 | -0.66 | -0.66 | -0.57 | -0.62 |
| $[-\gamma/\eta]$ | ( 0.07 ) | ( 0.07 ) | ( 0.10 ) | ( 0.04 ) | ( 0.07 ) |
| ln(market share) | -2372 | -2684 | -2643 | 128 | -2395 |
| $[-(1-\sigma)/\eta]$ | ( 723 ) | ( 778 ) | ( 920 ) | ( 328 ) | ( 684 ) |
| ln(nest share) | -1807 | -1729 | -1739 | -1474 | -1795 |
| [-$\sigma/\eta$] | ( 655 ) | ( 697 ) | ( 696 ) | ( 381 ) | ( 664 ) |
| Observations | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 | 1,053,058 |
| $ja$ groups | 37,794 | 37,794 | 37,794 | 37,794 | 37,794 |
| Cragg-Donald stat | 29.1 | 14.2 | 9.9 | 41.0 | 16.5 |

Notes: Sample includes monthly observations Jan 1999 - Mar 2008 of all passenger cars and light trucks age 0-25. Model*age fixed effects, monthly time dummies, and model year dummies are included. Column (1) is the primary specification from Table 3, column (1). Nest share is the share of all vehicles in the same class. Columns (2)-(5) instrument for $G_{jat}$ with a value similar computed with some change. The instruments in columns (2) and (3) are computed with gas prices lagged by 1 month and 1 year respectively. Column (4) uses the average gallons per mile within a model*age. Column (5) assumes that all vehicles travel 12,000 miles per year. Standard errors are robust and clustered by $ja$ (model * age).

Table 13: **Welfare Effects of Behavioral Feebate**

| | |
|---|---|
| **Effects on New Vehicle Market** | |
| Feebate Pivot (MPG) | 19.0 |
| $\Delta$Quantity Above Pivot (%) | 31.0 |
| $\Delta$Quantity Below Pivot (%) | -49.8 |
| $\Delta$Average MPG | 2.36 |
| **Welfare Effects, Excluding CO2** | |
| $\Delta CS^c$: Change in Choice Consumer Surplus ($/person) | -17.0 |
| $\Delta CS^b$: Change in Internality ($/person) | -32.1 |
| $\Delta CS^h$: Change in Hedonic Consumer Surplus ($/person) | 15.1 |
| **Gasoline and CO2 Effects** | |
| $\Delta$Gasoline Use (gallons/person) | -37.5 |
| $\Delta$Gasoline Costs ($/person) | -82.1 |
| $\Delta$CO2 Emissions (metric tons/person) | -0.3 |
| $\Delta$Climate Damages ($NPV/person) | -7.6 |

"Notes: All welfare and CO2 numbers are net present values over the lifetime of the vehicle, for a counterfactual policy that affects one model year of sales. Marginal damages of CO2 are assumed to be $30 per metric ton. Change in choice consumer surplus also includes the recycled net revenues from the feebate policy."

# A Appendices (Not For Publication)

## A.1 Appendix 1: Dynamic Consumer Choice Model

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 $\mathcal{JA}$ of possible model-by-age combinations. Owning vehicle $ja$ at time $t$ forces expected one-period gasoline expenditures $\widetilde{G}_{jat}$ and gives one-period utility flow $\widetilde{\psi}_{ijat}$. This individual-specific utility flow is the sum of average utility $\widetilde{\psi}_{jat}$ and an individual taste error $\widetilde{\varepsilon}_{ijat}$. 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:

$$\begin{aligned} \max_{\mathcal{JA}} u_{ijat} = & -\eta p_{jat} - \gamma \widetilde{G}_{jat} + \widetilde{\psi}_{ijat} \\ & + \beta \max \left\{ \max_{\mathcal{KA}} \{u_{ikat+1}\} + \eta p_{j(a+1)(t+1)} - \lambda_{ja}, u_{ij,a+1,t+1} + \eta p_{j,a+1,t+1} \right\} \end{aligned} \quad (22)$$

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 $E[\mathcal{JA}_{t+s}] = \mathcal{JA}_t$

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:

$$u_{ijat} = -\eta p_{jat} + \sum_{s=0}^{\tau-1} \beta^s \left( \gamma \widetilde{G}_{j,a+s,t+s} + \widetilde{\psi}_{ij,a+s,t+s} \right) + \beta^\tau \left( \eta p_{j,a+\tau,t+\tau} - \lambda_{ja} \right) + \beta^\tau u_{ija(t+\tau)} \quad (23)$$

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:

$$p_{jat} = \frac{1}{\eta} \left[ -(\ln s_{jat} - \ln s_{0t}) - \gamma G_{jat} + \sigma \ln s_{(j/n)at} + \psi_{jat} \right] \quad (24)$$

Substituting this into the utility function, we have:

$$
\begin{aligned}
u_{ijat} &= -\eta p_{jat} + \sum_{s=0}^{\tau-1} \beta^s \left(\gamma \widetilde{G}_{j,a+s,t+s} + \widetilde{\psi}_{ij,a+s,t+s}\right) \\
&+\beta^\tau \left( -(\ln s_{j,a+\tau,t+\tau} - \ln s_{0,t+\tau}) + \sigma \ln s_{(j/n),a+\tau,t+\tau} + \sum_{s=0}^{L-1-a-\tau} \beta^s \left(-\gamma \widetilde{G}_{j,a+\tau+s,t+\tau+s} + \widetilde{\psi}_{j,a+\tau+s,t+\tau+s}\right) - \lambda_{ja} \right) \\
&+ \beta^\tau u^*_{ijat+\tau} \\
&= -\eta p_{jat} + \sum_{s=0}^{L-1-a} \beta^s \left(\gamma \widetilde{G}_{j,a+s,t+s} + \widetilde{\psi}_{j,a+s,t+s}\right) + \sum_{s=0}^{\tau-1} \widetilde{\varepsilon}_{ij,a+s,t+s} \\
&+ \beta^\tau \left(-(\ln s_{j,a+\tau,t+\tau} - \ln s_{0,t+\tau}) + \sigma \ln s_{(j/n),a+\tau,t+\tau} - \lambda_{ja}\right) + \beta^\tau u_{ijat+\tau} \qquad (25)
\end{aligned}
$$

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:

$$
u_{ijat} = \eta(w - p_{jat}) - \gamma G_{jat} + \psi_{jat} + \epsilon_{ijat} \qquad (26)
$$

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

$$
\begin{aligned}
G_{jat} &= \sum_{s=0}^{L-1} \beta^s \widetilde{G}_{j,a+s,t+s} \qquad (27) \\
\psi_{jat} &= \sum_{s=0}^{L-1} \beta^s \widetilde{\psi}_{j,a+s,t+s} + \beta^\tau \left(-(\ln s_{j,a+\tau,t+\tau} - \ln s_{0t+\tau}) + \sigma \ln s_{(j/n)at} - \lambda_{ja}\right) \\
\varepsilon_{ijat} &= \sum_{s=0}^{\tau-1} \beta^s \widetilde{\varepsilon}_{ij,a+s,t+s}
\end{aligned}
$$

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.

For the dynamic decision to simplify directly to the static model, the critical assumption is therefore that consumers believe that the market is stationary. For our estimator based on the static model to be *unbiased* given the true dynamic nature of consumers' decisions is of course a weaker requirement than that the models be *identical*. In principle, simplifying to the static model can introduce any additional error, as long as that error is uncorrelated with $\Delta G_{jat}$. This means that the critical assumption from above, that consumers believe that the market is stationary, could in principle be weakened. For our estimator to be unbiased, consumers need not believe that the market is stationary, but they must believe that any "non-stationarities" - changes in prices and characteristics of future products - are uncorrelated with changes in gasoline prices. The implications of this assumption are discussed in the text, and in particular, section 7 provides evidence that our results are not sensitive to this assumption.

## A.2 Appendix 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.

### A.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 can be matched to 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. A single 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 a sale open to all buyers, in a physical auction. These predicted values are then exponentiated to obtain monthly price estimates for that model and model year.

The JD Power dataset consists of monthly summaries of individual dealer-to-customer transactions by vehicle (at the level of a VIN prefix) and transaction type (cash, lease, or financed). Mean monthly prices are adjusted for customer rebates and any difference between the negotiated trade-in price and the trade-in vehicle's actual resale value.

Mean new and used vehicle prices for selected model years are shown in Appendix Figure A1. 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 since all regressions are run with model by age fixed effects.

All regressions are weighted by the number of observations in the price data sets. This assigns higher weight to vehicles for which a more precise estimate of price is available, and a smaller weight to "exotic" vehicles that may vary substantially in price for reasons other than gas costs. However, since prices are taken from two different data sets, we scale the weights on new vehicles so that the mean weight for new vehicles and one year old vehicles is equal. Our results are qualitatively similar without this reweighting, and when new vehicles are excluded from the analysis.

**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. Total registered quantities for selected model years are shown in Appendix Figure A2.

### A.2.2 Vehicle Attributes

**Fuel Economy** Since 1975, the EPA has employed a consistent test, called a dynamometer test, to measure fuel economy.[29] 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. To construct a Composite fuel economy rating from the reported MPG's between 1985 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.

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:

$$\begin{aligned}\text{New Adjusted City} &= 1/(0.003259 + 1.1805/\text{Lab City}) \\ \text{New Adjusted Highway} &= 1/(0.001376 + 1.3466/\text{Lab Highway})\end{aligned} \tag{28}$$

To construct the revised Composite rating, the EPA changed these weights for the harmonic mean to 57% and 43% for City and Highway, respectively. In recent years, these revised ratings were the ones made public to consumers. The EPA also retroactively changed its fuel economy ratings for old vehicles, now assuming that the changes between original and new adjustments 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. For 1985-2008, we use the (retroactively phased in) New Adjusted EPA methodology in our primary specification, to reflect analysts' best guess at the true value of MPG. Using the alternative construction of MPG - consumers' best guess - does not statistically or substantively change the results, giving $\widehat{\frac{\gamma}{\eta}} = 0.63$.

Greene, *et al.* (2007) report that fuel economy in used cars degrades at an average of 0.07 MPG per year. We further adjust both measures of MPG to account for this. The distribution of fuel economy constructed for our primary specification (in miles per gallon) is shown for selected model years in Appendix Figure A3.

**Other Attributes** In selected specifications, we use data on vehicle characteristics, including horsepower, weight, wheelbase, torque, ABS brakes, traction control, and stability control. 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.

[29] The information in this and the following paragraph is from the U.S. Environmental Protection Agency (2008). We acquried the EPA adjusted and unadjusted test data for 1975-2008 directly from researchers at Oak Ridge National Laboratory.

### A.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.

Note that a vehicle of a given model year typically begins being sold in September of the previous calendar year. The gasoline prices used to construct the instrument $G_{j0(t-a)}$ were thus the September-August mean gasoline prices.

**Gasoline Prices** Our source of gasoline price data is the US Energy Information Administration (EIA). We use US City Average Motor Gasoline Retail Price for all types of gasoline, which are available on a monthly basis from Table 9.4 of the EIA's Monthly Energy Review.

**Oil Futures** We acquired the entire history of futures prices for Light Sweet Crude Oil (LSCO) on the NYMEX and Intercontinental Exchange. Oil futures prices are transformed into gasoline price expectations in constant 2005 dollars using the following approach.

First, futures prices are denominated in future dollars at the delivery date, meaning that they must be discounted to current dollars using a measure of inflation expectation. Our measure of inflation expectations is from the difference in yield rates between standard and inflation-protected (TIPS) treasury bonds, available from http://www.federalreserve.gov/releases/h15/data.htm#fn14. We use the five year bonds for futures with maturities zero to six years in the future, the seven year bonds for maturities six to eight years in the future, and the ten year bonds for maturities more than eight years in the future. Before the TIPS bonds were sold in 2003, we use inflation expectations of 2%. This is approximately consistent with actual observed inflation over the period 1998-2002.

Second, all data are deflated into real July 2005 dollars. Third, these constant-dollar oil price expectations are converted into gasoline price expectations. This is done using the slope and intercept coefficient estimates from a linear regression of historical national city average retail gasoline prices on spot LSCO prices. Among other things, this assumes that refiners' margins will be constant over time.

Finally, to model expectations for periods beyond the last futures contract settlement date observed at each time $t$, we use a simple model of mean-reverting expectations, where deviations at time $t$ from a mean of $1.50 per gallon 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 $\rho$ is 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 estimation gives $\widehat{\rho} = -0.057$, meaning that the market expected recent gasoline price increases to decay back to $1.50 per gallon at 5.7 percent per year.

**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 almost exactly equivalent to real July 2005 dollars.

**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 next year, $q_{ja(t+1)}$, out of the number of vehicles registered today, $q_{jat}$, 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. The estimation coefficients are used to predict a series of probabilities that each vehicle in the data set survives to time $t+1$ conditional on surviving to time $t$, for current and future values of $t$. These are multiplied to compute probabilities that each vehicle survives an additional $s$ years beyond its current age, for all positive $s$ up to $L = 25$. 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. Appendix Figure A4 illustrates these data by showing average annualized VMT as a function of vehicle age.

**Accounting for Elasticity on the Intensive Margin** While our primary specification of $G_{jat}$ assumes that the elasticity of vehicle miles traveled with respect to gasoline prices is negligible, we also present an alternative specification that accounts for this elasticity. We now detail how this alternative specification was derived. The model must capture two effects. 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 adopt estimates of short run elasticity of gasoline demand from three recent papers. Hughes, Knittel, and Sperling (2007) find that between 2001 and 2006, the short run elasticity was between -.034 and -.077. Small and Van Dender (2007) find that with covariates at their 1997-2001 levels (the latest years in their study period), the short run elasticity is -.022. Davis and Kilian (2009) use an instrumental variables procedure identified off of state-level changes in gasoline taxes and estimate a short run elasticity of -0.46. We assume an elasticity of -0.5 to conservatively overstate the potential effects of intensive margin elasticity.

We generate expected VMT at any possible gasoline price, using the following constant elasticity formula:

$$\log(m_{jas}) = (-0.5)\cdot\log(g_s) + \kappa_0 \tag{29a}$$

$$\kappa_0 = \log(m_{ja,2001}) - (-0.5)\cdot\log(g_{2001}) \tag{29b}$$

Appendix Figure A5 shows a vehicle owner's demand for VMT. In 2001, gas prices are $g_{2001}$ and consumers 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 $m_{jas}(g_s)$. The new annual gasoline cost is now the unshaded rectangle bounded by $g_s/f_{jas}$ and $m_{jas}$. The values of $G_{jat}$ in our alternative specification are calculated from these adjusted VMT values $m_{jas}$.

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 in Appendix Figure A5. Summing this over the future years of the vehicle's life, we have an adjustment denoted $I_{jat}$. For simplicity, we assume a linear demand curve between $m_{ja,2001}$ and $m_{jas}$ in computing the utility change:

$$I_{jat} = \sum_{s=t+1}^{t+(L-1-a)} -f_{jas}^{-1} \cdot (g_{2001} + g_s) \cdot (m_{ja,2001} - m_{jas}) \cdot \phi_{jas} \cdot \beta^{s-t} \tag{30}$$

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. In our alternative specification that accounts for intensive margin elasticity, we move $I_{jat}$ to the left hand side and estimate equation (9) with $p_{jat} + I_{jat}$ as the dependent variable.

### A.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 1999 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. We drop vehicles which do not use gasoline or are not part of the set of substitutable passenger vehicles, such as delivery vehicles and motor homes. Finally, because we use fixed effects, a small number of $ja$ vehicles with prices observed only at one $t$ must be dropped from the estimation. Appendix Figure A6 shows the data coverage.

## A.3 Appendix 3: Predicted Substitution Patterns

This appendix presents a "reality check" of the substitution patterns predicted by our parameter estimates in the nested logit model. Table 14 presents example substitution patterns between new vehicles in the 2007 choice set predicted by the nested logit model for the two most popular vehicles in the compact, SUV, and pickup vehicle classes. The table shows the own-price and cross-price elasticities computed by increasing the price of all submodels within a given new model by five percent. Instead of using the analytical nested logit cross-price elasticities, these are generated while holding constant the quantities of all used vehicles and of the outside option, using a modification of the Berry, Levinsohn, and Pakes (1995) contraction mapping procedure.[30]

[30] The modified contraction is:

$$\widetilde{p}'_{jat} = \widetilde{p}_{jat} - \frac{1}{\eta}\left(\ln s_{jat} - \ln \widehat{s}_{jat}(\widehat{\eta}, \widehat{\gamma}, \widehat{\psi}, \widehat{\sigma}, \widetilde{p}_{jat}, F_{jat})\right) \tag{31}$$

In this equation, $\widetilde{p}'$ indicates the used vehicle prices to be used in the subsequent iteration of the contraction mapping. Note that if the prices of all substitutes are unchanged, the procedure would immediately solve for $\widetilde{p}_{jat} = p_{jat}$, the intially-observed equilibrium used vehicle prices, as the $\widehat{\psi}_{jat}$ were themselves implied by the initial market

The table shows, for example, that the Honda Civic had 325 thousand new model sales in 2007. If its price were increased by 5 percent, quantity demanded would decrease by $6.15 \times 0.05 = 31$ percent. Quantity demanded of the Toyota Corolla would increase by $0.18 \times 0.05 = 0.9$ percent. This table brings to light the substitution patterns implied by the nested logit specification. If the price of a vehicle increases, the quantities of all vehicles in other nests change by essentially the same percent, after accounting for small changes in used vehicle prices. Cross-price elasticities are higher within the same nest, depending on the value of the substitution parameters $\sigma$, but are effectively the same for all substitutes within the nest.

We can similarly compute the aggregate price elasticity of demand in the new vehicle market by increasing the prices of all new vehicles by 10 percent and resimulating quantities, holding constant the market shares of each used vehicle. Predicted 2007 new vehicle sales drop from 14.16 million to 7.23 million. This secant calculation gives an overall new vehicle market price elasticity of -4.89.

An additional approach to testing the sensibility of these estimated substitution parameters is to back out the implied markups that auto manufacturers would be applying to their vehicles if they are playing a static Nash multi-product pricing game.[31] As shown in Table 14, the implied markups for our six example new vehicles range from 9.3 percent to 18.4 percent, or around \$2600 per vehicle.

If anything, the estimated $\widehat{\eta}$ and $\widehat{\sigma}$ are high and low, respectively: relative to our priors, consumers are estimated to be slightly more price elastic and somewhat more willing to substitute across vehicle classes. To test whether this affects the results, we experimented with alternative specifications that fixed lower values of $\eta$ and higher values of $\sigma$ in the second stage of equation (9). Both of these types of changes produce estimates of $\frac{\gamma}{\eta}$ that are further away from unity.

---

equilibrium with no feebate.

This equation can be derived from the original BLP contraction mapping by subsituting in for their average utility $\delta_{jat}$:

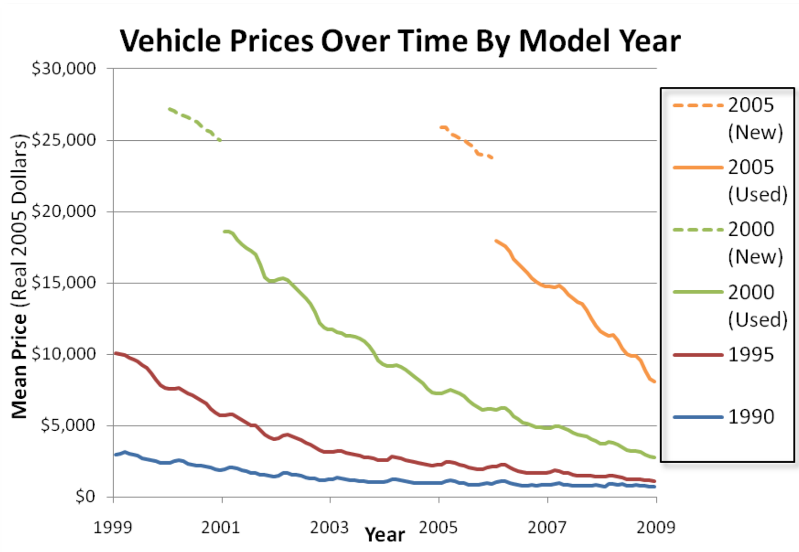
$$\delta_{jat} = -\eta p_{jat} - \gamma G_{jat} + \psi_{jat}$$

Given that $G_{jat}$ and $\psi_{jat}$ are constant throughout the procedure, solving for a set of $\delta$ values that equate observed with predicted shares is equivalent to solving for a set of equilibrium used vehicle prices.

[31] See Berry, Levinsohn, and Pakes (1995) or Petrin (2002) for details on computing the markups implied by estimated substitution patterns in the static Nash pricing game.

# Appendix Figures

## Figure A1: Vehicle Prices




## Figure A2: Vehicle Quantities

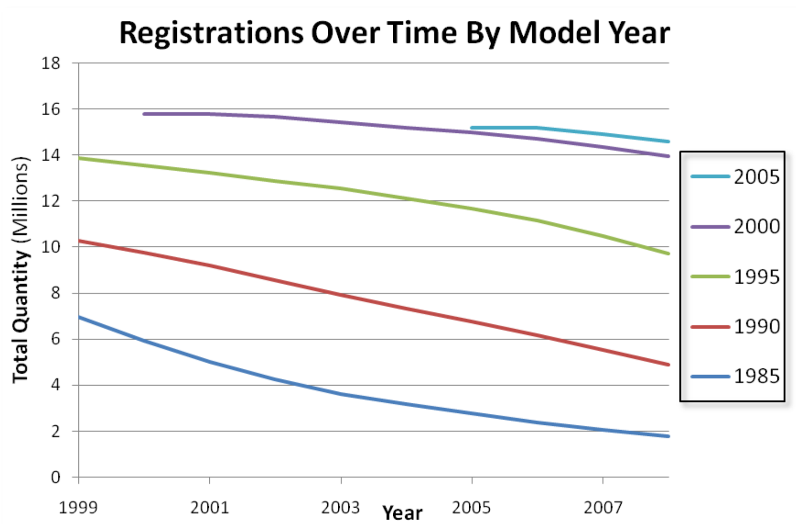


Figure A3: Fuel Economy Ratings of Vehicles Registered in 2007

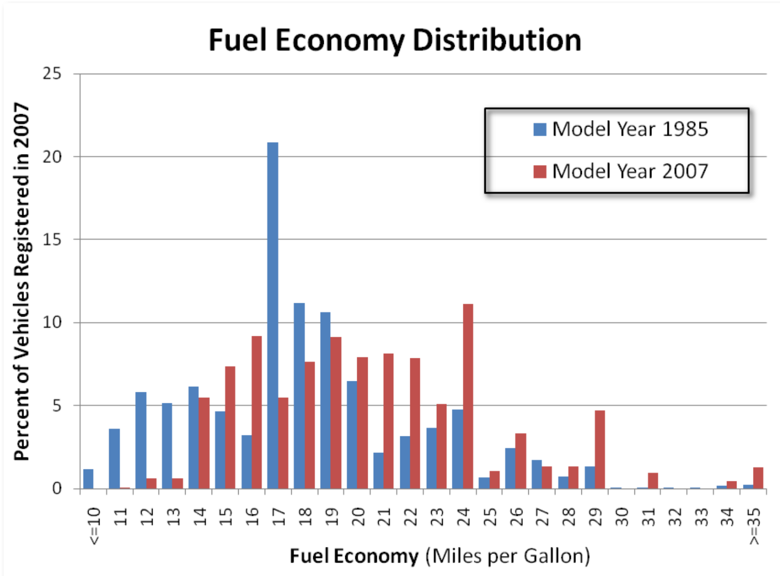



Figure A4: Vehicle Miles Traveled By Vehicle Age

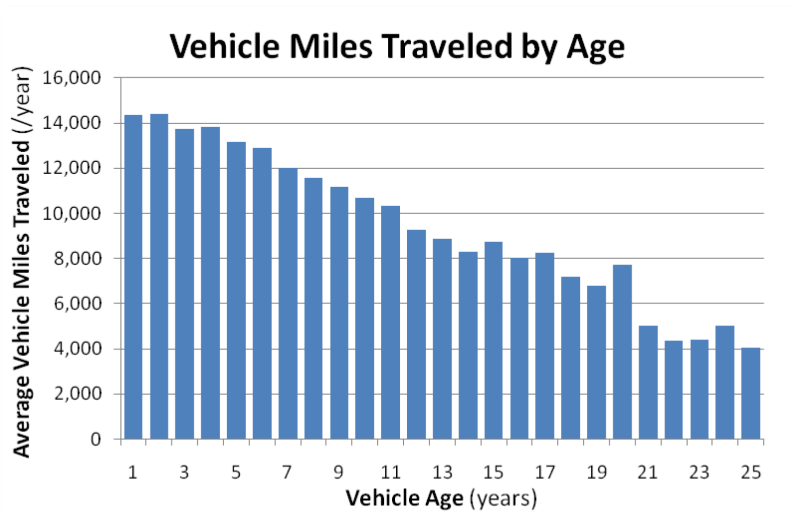


**Figure A5: Intensive Margin**

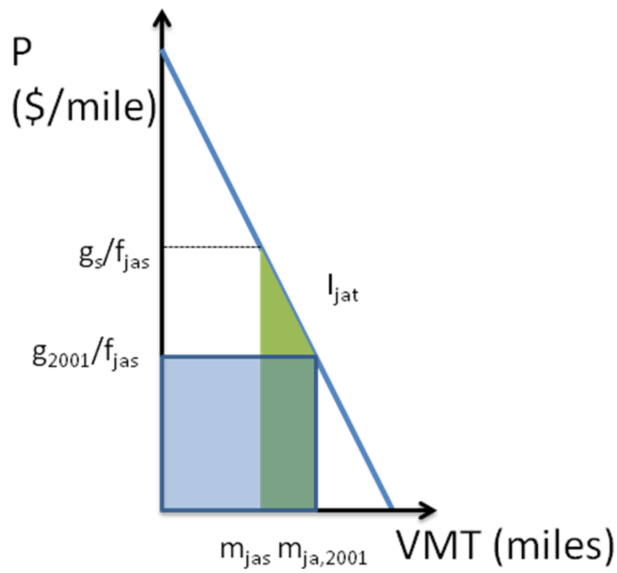
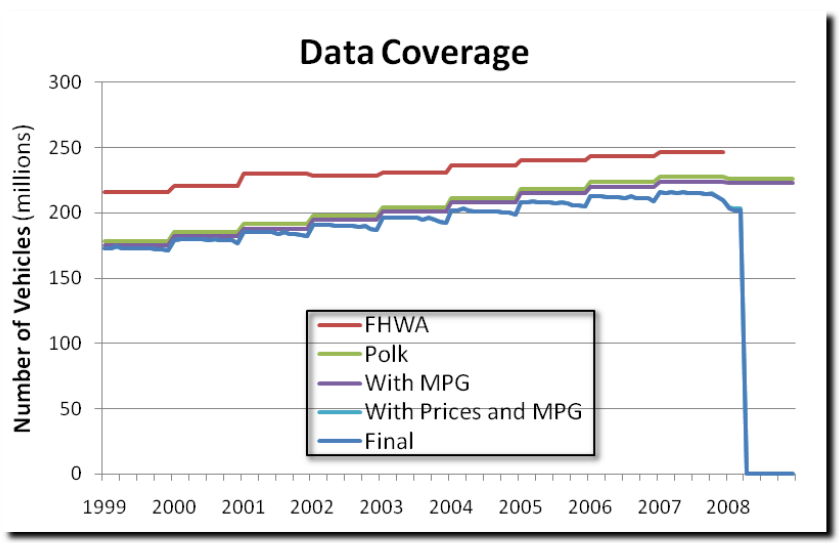



**Figure A6: Data Coverage**



## Appendix Table

Table 14: **Predicted New Vehicle Own and Cross-Price Elasticities**

| **Vehicle** | Quantity | Civic | Corolla | Equinox | F-Series | Pathfinder | Ram |
|---|---|---|---|---|---|---|---|
| Civic | 325 | -6.15 | 0.14 | 0.03 | 0.20 | 0.02 | 0.09 |
| Corolla | 300 | 0.18 | -5.30 | 0.03 | 0.20 | 0.02 | 0.09 |
| Equinox | 77 | 0.13 | 0.11 | -7.09 | 0.20 | 0.03 | 0.09 |
| F-Series | 360 | 0.14 | 0.11 | 0.03 | -8.08 | 0.02 | 0.12 |
| Pathfinder | 38 | 0.13 | 0.11 | 0.04 | 0.20 | -8.73 | 0.09 |
| Ram | 173 | 0.14 | 0.11 | 0.03 | 0.26 | 0.02 | -7.81 |
| Markup (%) | | 14.6 | 18.4 | 12.8 | 12.3 | 9.3 | 10.6 |

Notes: Values shown are the elasticity of demand for the vehicle in the row name with respect to the price of the vehicle in the column name. To generate this table, the price of all submodels of the given model were increased by five percent, and the market shares of all other new vehicles were re-simulated. The simulation adjusts the prices of individual used vehicles and the price level of all new vehicles so as to hold constant the market shares of used vehicles and of the outside option. Quantities are 2007 new model sales in 1,000s. Markups are those implied by a static Nash multiproduct pricing game.