

Gas Prices and Endogenous Product Selection in the U.S. Automobile Industry

Jacob Gramlich*
Yale University

Abstract

I develop and estimate a model of the U.S. automobile industry in which firms choose the fuel efficiency of their new vehicles. I then use the model to analyze the 2008 gas price increase of 28%. Firms face a technological frontier between providing fuel efficiency and other quality, and the gas price shifts incentives to locate along this frontier. Where firms locate along this frontier is of environmental and policy significance. The demand model is nested logit, and supply is differentiated products oligopoly. The model is estimated using data from the US automobile market from 1971-2007. The model predicts a 2008 sales decline of 11.9% and sales-weighted fuel efficiency increase of 27.9%. These are close to actual figures. I then demonstrate that firms' have incentives to raise fuel efficiency approximately 20% for vehicle offerings in 2009 and beyond. Contributions to previous work include modeling product choice, relaxing restrictive identifying assumptions, and obtaining more realistic estimates of fuel efficiency preference. The model can be used to analyze various policy and environmental changes.

KEYWORDS: Automobiles, endogenous product choice, environmental policy

J.E.L. CLASSIFICATION: D21, D12, H23, L1, Q38

Job Market Paper

November 20, 2008

Comments Welcome

*Dept. of Economics, Yale University, 28 Hillhouse Avenue, New Haven, CT 06511. jacob.gramlich@yale.edu
I thank my advisers Steven Berry and Philip Haile for guidance. I also thank Justine Hastings, Patrick Bayer, Philip Keenan of General Motors, Martin Zimmerman, William Brainard, Costas Arkolakis, Larry Samuelson, Alvin Klevorick, Yuichi Kitamura, Donald Andrews, Robert Shiller, Joshua Lustig, Christopher Conlon, Myrto Kalouptsidi, and participants in the Yale microeconomics prospectus workshop for helpful comments. Steven Berry, Jim Levinsohn, and Ariel Pakes generously shared data for this research. The Cowles Foundation generously purchased data for this research.

1 Introduction

This paper investigates how changes in the economic and policy environment affect firms' decisions about product characteristics. In particular I investigate the automobile industry, which has recently seen dramatic changes in both gas prices and regulation. In 2008 gas prices increased 23% through August.¹ Also in 2008, the federal government increased fuel efficiency regulations (CAFE Standards) for the first time in over 30 years.² These changes affect the fuel efficiency that auto manufacturers place in their new automobiles.³ The question this paper seeks to address is this: How do economic and policy changes affect firms' choices of fuel efficiency in their new vehicles? Will firms change fuel efficiencies, and if so, by how much?

The industry for new automobiles is large and environmentally significant. 2007 revenues exceeded \$400 billion, which is 3% of US GDP. Passenger vehicles consume 20% of our nation's energy⁴ and emit 20% of our nation's CO₂.⁵ Growing concerns over global warming, energy dependence, and national security have led to increased focus on energy policy. Automotive fuel efficiency is one of the most direct measures of a vehicle's environmental impact, so fuel efficiency itself has become a target of energy policy. Understanding how firms choose product characteristics in this industry is therefore crucial to informing these policies.

Automobile manufacturers can increase fuel efficiency in their vehicle fleet in three ways. First, because they are multi-product firms, they may introduce more fuel efficient automobiles and/or discontinue less efficient models. Second, they may increase fuel efficiency of an existing vehicle, holding other quality constant. Third, they may increase fuel efficiency by trading off quality of the vehicle. For instance, one may achieve a fuel efficiency increase by decreasing engine power or vehicle weight.⁶ This paper focuses on the third mechanism. In the "medium run" (3-5 years) the first two mechanisms are limited because they change the automobile's "class" or "segment."⁷

¹In autumn 2008, at the time of writing, they are falling.

²CAFE Standards (Corporate Average Fuel Economy Standards) stipulate minimum sales-weighted fuel efficiencies for automobiles sold in America, and carry fines for non-complying firms. Until January 2008, the standards had been in place without major revision since 1975 when they were first instituted.

³Fuel efficiency is the rate of gasoline consumption. In this analysis it is miles per gallon.

⁴NEED (2008).

⁵EPA (2008).

⁶Power provides performances. Weight provides luxury and/or safety.

⁷For example, SUV, or Compact Car.

These are larger changes and take longer to implement. In this analysis I hold vehicle segments constant and focus on the tradeoff between quality and efficiency within vehicle segment. Changes in the gas price induce changes in firms' optimal locations along this efficiency v. quality frontier. Firms face an exogenously changing gas price, and consumers are more sensitive to fuel efficiency when gas prices are high. High gas prices give firms the incentive to provide more fuel efficiency and less quality, while low gas prices provide the opposite incentives.⁸

Empirical models of industry, including seminal work on automobiles (Berry, Levinsohn, and Pakes, 1995), have focused on correcting for the endogeneity of price.⁹ However, a limitation of these models has been the treatment of product characteristics. The limitation has been twofold. First, characteristics have been assumed to be fixed exogenously. This is limiting when changes in these characteristics (such as fuel efficiency) are of direct interest. Second, in contrast to prices, the models have not allowed characteristics to be correlated with the unobserved cost and quality shocks. In fact, the identifying assumption of these models has been strict orthogonality between characteristics and the unobserved shocks. This restriction is implausible for the same reason it is with prices: prices and characteristics are both choice variables of the firm, and are chosen in response to an observed economic environment which includes the shocks. So while the orthogonality assumption is useful in estimation, it is implausible in many industries including automobiles.¹⁰

My model addresses both of these limitations in the context of automobiles. First, as mentioned, I develop a model which allows firms to choose product characteristics. This allows me to analyze changes in fuel efficiency choices themselves. Second, I relax the restrictive identifying assumption commonly used to estimate demand. I use alternative estimation moments based on the timing of events in the auto industry. In particular, I note that firms commit to product characteristics before the year of sale. Characteristics decisions are made to maximize expected profit, where expectations

⁸In the wake of the 1979 gas price increase, GM President Elliott M. Estes noted that consumer demand placed pressure on firms to offer more fuel efficient choices. "Consumers seem to be reasserting themselves as our No.1 taskmaster".

⁹Price is correlated with unobserved cost and quality shocks, and allowing for this correlation improves model fit and predictions. "Unobserved" means unobserved to the econometrician, but observed to market participants.

¹⁰A separate but related issue is that in addition to observing the shocks, firms might actually choose these shocks or at least a portion of them. In the automotive industry, firms likely choose a portion of these shocks. I do not allow for this in my model, however, because doing so would require additional assumptions in order to split a single unobservable quantity into two unobservable portions (chosen and unchosen).

are taken over the gas price. After characteristics decisions are made, the gas price changes, potentially causing ex post regret in characteristic choice. The timing of the model implies that this regret is not foreseeable, systematic, or predictable. In other words, the regret is uncorrelated with anything known at the time of the decision (Hansen and Singleton, 1982). This model implication becomes an estimation moment and allows me to relax the implausible assumption that unobserved qualities are orthogonal to observed characteristics. Given that I allow free correlation between product characteristics and unobserved qualities, I report these correlations in the estimation results and show them to be quantitatively significant.

A third contribution of the model relates to the representation of consumer preferences for fuel efficiency. Previous models, including seminal work,¹¹ have had difficulty finding parameter estimates that show consumers care about fuel efficiency. Parameters on preference for fuel efficiency have been biased towards zero. The reason for this is that in automobiles, fuel efficiency (*mpg*) is negatively correlated with other characteristics that provide utility. Some of these characteristics are observed and easily controlled for, such as horsepower and weight. Others, however, are not. Engine characteristics such as timing of transmissions gear ratios can also affect the fuel efficiency v. performance tradeoff but are harder to capture in data. I correct for this aspect of unobserved product quality by controlling for both the economic and quality effects of fuel efficiency in consumer preferences. Unlike previous work, my resulting parameter estimates indicate that consumers care considerably about fuel efficiency.¹² There are quality tradeoffs in providing fuel efficiency, even beyond the characteristics most commonly observed and controlled for in automotive data.

I use the model to analyze the 2008 gas price increase. To check the predictive ability of the model, I compare out-of-sample predictions to 2008 actual figures. The model matches sales composition changes well. Actual aggregate sales are down 12% through August from 2007 levels, compared to a prediction of 11.9%. The model also predicts an increase in sales-weighted fuel efficiency of 28%, a number consistent with the large 2008 reductions in purchases of SUVs and light trucks. I then demonstrate firms' incentives to raise fuel efficiency approximately 20% in subsequent vehicle offerings (2009 and beyond). The model can be used to analyze various changes in the market

¹¹Berry, Levinsohn, and Pakes (1995).

¹²I discuss the magnitude and distribution of these preferences further in the estimation results (section 6).

and policy landscape, including gas price changes, gas taxes, and changes in CAFE standards.

2 Relation to Literature

This paper is related to literatures on both the automotive industry and endogenous product selection. The automotive literature has recently focused on endogeneity of prices [Berry (1994), Berry, Levinsohn, and Pakes (1995)], as well as various policy questions [Goldberg (1995), Gruenspecht (1982), Berry, Levinsohn, and Pakes (1999), Kleit (1990) and numerous others]. This paper investigates a similar question to Pakes, Berry, and Levinsohn (1993) (the effects of a gas price increase on the automobile market) but adds a model of product choices in order to analyze fuel efficiency change. The model of product choice is new to the automobile literature.¹³

This work is also related to a growing literature on endogenous product selection. This literature reflects the importance of changes to product characteristics, not just pricing changes. The industries and applications within which this question have been studied are various, and the modeling obstacles and adaptations are as varied. Some examples are early work in this area Mazzeo (2002) which studied binary and trinary entry decisions in the motel market. Lustig (2008) uses cross sectional variation in market structure to investigate health insurance quality. Sweeting (2007) uses a dynamic framework to analyze the radio industry, and Crawford and Shum (2007) use the theory of monopoly screening to study the cable industry.

¹³An exception is Goldberg (1998) which models the domestic/foreign production location decision in response to CAFE standards.

3 Model

3.1 Demand

Consumer demand feeds into firms' profit functions. Each year, U.S. households take as exogenous both the gas price and the product offerings of firms. They choose a new automobile (or no new automobile - the outside good) to maximize a conditional indirect utility function:

$$u_{ijt} = u(p_j, econ_j, qual_j, X_{jt}, \xi_j, \epsilon_{ijt}) \quad (1)$$

A few comments about notation, both for this equation and throughout paper, are in order. The subscripts are i for individual, j for automobile model, and t for time (year). Note that any j subscript could itself be subscripted by t because model attributes change from year to year, but I suppress this notation for ease of exposition. Where t subscripts do appear, they are to emphasize that some variables are common across all j in a given year (such as macroeconomic variables in X_{jt}) or to emphasize the frequency of some errors or decisions (such as $\tilde{\epsilon}_{ijt}$). I use the terms “model,” “product,” and “vehicle” (subscript j) interchangeably. Vehicle “type” (subscript v), vehicle “segment” (subscript s), and vehicle “sub-segment” (subscript ss) refer to specific levels of auto classification that come in my data. These are displayed in Figure 2. Note that t would be the natural subscript for vehicle “type,” but it already subscripts time. To avoid confusion I use v instead. Throughout the discussion, scalar quantities (such as p_j) will be plain text, while vector quantities (such as \mathbf{p}) will be in bold.

The utility specification in equation (1) indicates that within an automotive sub-segment, consumers have preferences over price (p), fuel economy ($econ$), and “other quality” $qual$. Other quality includes things such as power, weight, acceleration, electronics, sportiness, interior room, etc. - in short, the collection of other vehicle attributes that must be traded off with fuel efficiency. X contains terms controlling

for sub-segment utility intercepts, segment-specific preference for fuel economy, and macroeconomic variables such as GDP growth. The macroeconomic variables are intended to capture the utility of the outside good, not purchasing a new automobile. ξ is unobservable quality not captured in the other utility covariates. $\tilde{\epsilon}$ is a nested logit error term.

Fuel efficiency (*mpg*) affects both the fuel economy (*econ*) and “other quality” (*qual*) of a vehicle through the technological tradeoff. The measure I use for *econ* is dollars-per-mile (*dpm*), which is equal to the gas price over the fuel efficiency ($\frac{p_{gas}}{mpg}$).¹⁴ This is a key component of the economic cost to operating a vehicle.¹⁵ This form implies the natural result that the marginal benefit of fuel efficiency, $\frac{\partial u}{\partial mpg}$, is increasing in the price of gas.¹⁶ The measure I use for *qual* is *mpg* itself. This may seem unusual but, conditional on the economic effects of fuel efficiency (*dpm*), higher *mpg* is strongly associated with lower “other quality.” This tradeoff can be pictured with the hypothetical isocost curve (for a manufacturer within a given sub-segment) that I propose. (Figure 1, in the text).

The frontier reflects tradeoffs between efficiency and, among other attributes, power (performance) and weight (luxury/safety). This tradeoff has been documented in previous work (Kleit, 1990). It is evident in my data as well. Figure 3 shows a picture for 2007 of the relationship between $\ln(mpg)$ and $\ln(hp)$ ¹⁷ as well as between $\ln(mpg)$ and $\ln(weight)$. This relationship is negative for any level of aggregation. Likewise, Table 2 shows the result in the form of a regression. $\ln(mpg)$ is regressed on $\ln(hp)$, $\ln(weight)$, and a time trend to capture advancing technology. This also shows a strong negative relationship between fuel efficiency and both power and weight. In my model I assume this tradeoff holds exactly within a sub-segment. The regression numbers in the Table are pooled over all vehicle types but, at finer levels of aggregation including the sub-segment, the results change little. If the strength of the relationship were constant, then the R-squared should drop at finer levels of aggregation due to the lower number of observations. However, few sub-segments drop below an R-squared of .4, most stay between .6-.8, and some even rise slightly (up to

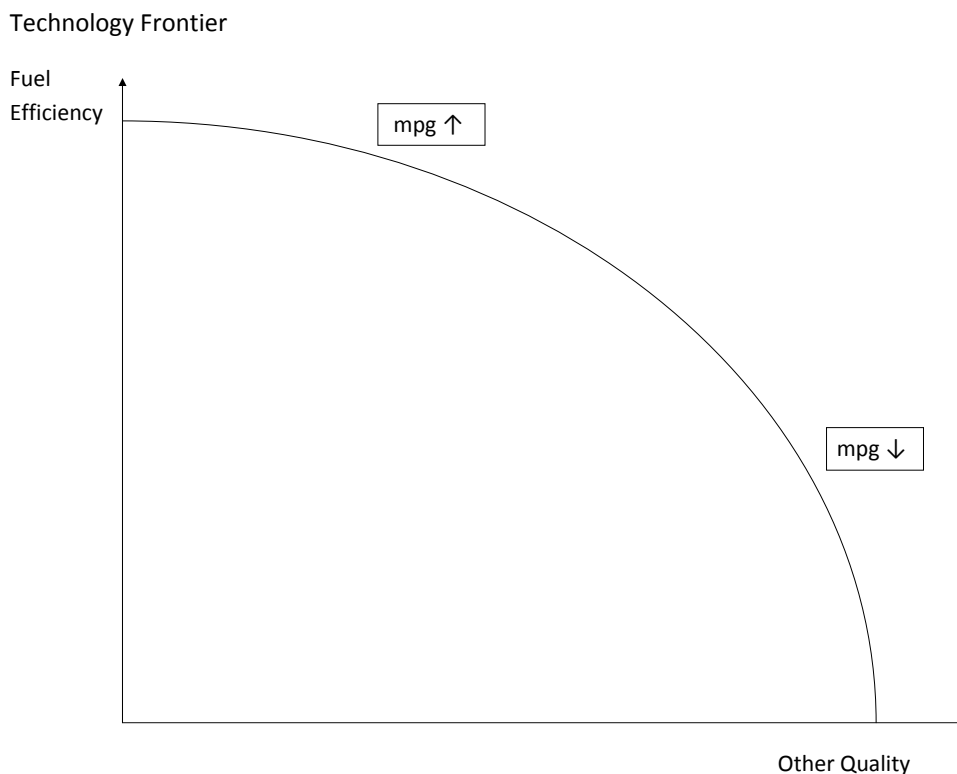
¹⁴Note the distinction between the technological parameter fuel **efficiency** (*mpg*) and the consumer preference parameter fuel **economy** (*dpm*).

¹⁵I do not separately model consumers’ vehicle utilization decisions.

¹⁶I.e., $\frac{\partial^2 u}{\partial mpg \partial p_{gas}} > 0$.

¹⁷*hp* is horsepower, a measure of engine power.

Figure 1: Within any automobile sub-segment:



.85) due to the increasingly specific technological relationship within a sub-segment.

There are other attributes of quality besides power and weight that are negatively correlated with *mpg*. Some of these are less easily quantified in the data. Engine characteristics such as timing of transmissions, gear ratios, turbo charging, cylinder shutdown, and variable valve timing can also affect the fuel efficiency v. performance tradeoff. The presence of these components of unobserved product quality is evidenced by demand parameter estimates. Conditional on fuel economy (*dpm*), *mpg* enters utility negatively, statistically significantly, and extremely robustly to various specifications. This result is so robust that I am unable to find a specification that reverses it, even controlling for every datum in my dataset. I have found no other piece of data that proxies for these quality compromises quite as well as *mpg* itself, so I use *mpg* as my proxy for other quality.¹⁸ My results indicate that in future auto de-

¹⁸I could alternatively estimate a hedonic index of other quality based on power, weight, and other things observed in the data. However, since I am not going to allow firms to choose these individual pieces, I simply use *mpg* itself which firms do choose.

mand estimation, even when including many utility covariates, the inclusion of both dpm and mpg helps to control for the unobserved quality implications of providing fuel efficiency.

Given the above definitions, and with an additional linearity and separability assumption, consumer utility in equation (1) becomes:

$$u_{ijt} = \alpha p_j + \beta_d dpm_j + \beta_m mpg_j + \xi_j + BX_{jt} + \tilde{\epsilon}_{ijt} \quad (2)$$

Note that dpm and mpg in fact measure the negative of the characteristics for which they proxy (fuel economy and other quality, respectively).¹⁹ The nesting structure designations in Figure 2 come from Wards Automotive. They are a 3-level structure, where each level is an increasingly specific level of automobile classification. The top level, vehicle type (v), are Cars, Utility Vehicles, and Trucks/Vans. The next level, segment (s), consists of designations such as Small Car and Sport Utility Vehicle (SUV). The third and most specific level, sub-segment (ss) designates categories such as Lower Small Car, or Luxury SUV. Finally, the model j are familiar names such as Honda Accord or Ford Explorer.

The nested logit error term $\tilde{\epsilon}$ contains parameters to be estimated ($\sigma_{ss}, \sigma_s, \sigma_v$) which describe the similarity of choices within a nest. To be consistent with utility maximization, parameter estimates should satisfy $1 \geq \sigma_{ss} \geq \sigma_s \geq \sigma_v \geq 0$ (Cardell, 1997). Appendix A contains details on the calculation of nested logit shares and σ 's.

Note that ξ_j from equation (2) will be used in estimation moments, although in a less restrictive way than in previous demand estimation.

¹⁹I prefer to let parameter estimates come out negative, rather than carry around a negative in the definition of the data.

3.2 Supply

The timing of the supply side model is depicted in Figure 4. One year before the year of sale of vehicles,²⁰ firms observe gas prices, cost shocks, and demand shocks. In response they commit to product characteristics for each vehicle, still one year in advance, to allow for a production lag. Then, in the interim, gas prices change. Finally, in the year of sale, firms choose a price for each vehicle and consumers make purchase decisions. The information sets known to all (H and I) as well as the ex post regret (R) will play a role in the estimation moments as discussed in section 5. Note that this timing allow firms to observe their cost and demand shocks when choosing their product characteristic, mpg .

Firms take actions to maximize expected profit. Firms' choices (of mpg and p) form a subgame-perfect Nash equilibrium. Any one firm f 's problem is:

$$\max_{\mathbf{mpg}_f} E_{p_{gas}} \left[\max_{\mathbf{p}_f} \Pi_f \right] \quad (3)$$

The specific form of firm f 's profits is:

$$\Pi_f = \sum_{j \in \mathfrak{S}_f} M_t [s_j(\mathbf{mpg}, \mathbf{p}; \theta)(p_j - mc_j(mpg_j; \theta)) - \lambda_{ft} \{bind_{ft}\} (CAFE_t - \overline{mpg_{ft}})] \quad (4)$$

Subscript f is for firm (recall that j is a model). \mathfrak{S}_f is the set of cars produced by firm f , as auto manufacturers are multiproduct firms. M_t is the size of the market (the number of U.S. households), s_j is the market share of model j , and is a function of \mathbf{mpg} and \mathbf{p} , the entire industry vectors. θ is a vector of demand and cost parameters to be estimated. $CAFE_t$ are the stipulated CAFE standards, and $\overline{mpg_{ft}}$ is the

²⁰As a robustness check I estimate the model assuming 3 and 5 year lags and discuss differences in section 7.

firm's average fuel efficiency for determining CAFE compliance.²¹ $\{bind_{ft}\}$ indicates whether the standards are binding for the firm, and λ_{ft} is a firm and time-specific shadow cost of violating the standards. Note that given the timing of the game, price p is a function of many things (including but not limited to mpg) in equilibrium, but this notation is suppressed for exposition. Recall that the t subscripts have also been suppressed in a number of places for ease of exposition.

Because of the Nash equilibrium assumption, derivatives of the profit function with respect to the choice variables lead to first order conditions of optimality. These hold with respect to the two choices (p_j and mpg_j) for each model:

$$s_j + \sum_{r \in \mathfrak{S}_f} \left[(p_r - mc_r) \frac{\partial s_r}{\partial p_j} + \lambda \{bind\} \frac{\partial \overline{mpg}_f}{\partial p_j} \right] = 0 \quad (5)$$

$$E_{p_{gas}} \left[\sum_{r \in \mathfrak{S}_f} \left[s_r \left(\frac{\partial p_r}{\partial mpg_j} - \frac{\partial mc_r}{\partial mpg_j} \right) + (p_r - mc_r) \frac{\partial s_r}{\partial mpg_j} + \lambda \{bind\} \frac{\partial \overline{mpg}_f}{\partial mpg_j} \right] \right] = 0 \quad (6)$$

Note that a single choice of mpg along the sub-segment frontier, or isocost curve, determines both the fuel efficiency ($dpm_j = \frac{p_{gas,t}}{mpg_j}$) and other quality (mpg_j) for consumer demand. Often a challenge in modeling product choice is finding exogenous variation to shift incentives to offer characteristics. Characteristics-based models of auto demand have tended to include around six preference characteristics of automobiles. Because choices of these attributes are interrelated (for example increasing fuel efficiency by decreasing engine power), allowing firms to choose any of these characteristics means the model must allow firms to choose all of them. Allowing firms to choose only some characteristics, while treating others as exogenous, would bias measured incentives. Having said this, however, it is difficult to find instruments to shift firms incentives to provide all six characteristics in the auto industry. Some industries have the benefit of having many local, geographic markets in the cross section that provide variation in market structure (e.g. Lustig (2008)). However, automo-

²¹CAFE standards calculate firm fuel efficiency as a harmonic average rather than a raw average. $\overline{mpg}_f = \frac{q_f}{\sum_{j \in \mathfrak{S}_f} \frac{q_j}{mpg_j}}$. Harmonic averages are used to average rates, as they indicate the rate of fuel consumption when all automobiles are driven the same distance. They are (mechanically) lower than raw averages, since low efficiencies pull down the average more than high efficiencies raise them.

bile supply is one, national market each year, so any shifter must have time-series variation.

My solution to this problem is the tradeoff between fuel efficiency and other quality within an automotive sub-segment. In the medium run (three to five years), it is plausible to assume that firms do not have time to switch cars from one segment to another.²² In addition, *within* a sub-segment, it is plausible to assume that consumers care only about the tradeoff between fuel efficiency and other quality. While all six characteristics might be relevant to consumers' sub-segment choices, within these sub-segments cars are relatively more similar on characteristics. It is plausible that the within-sub-segment comparisons happen along fewer dimensions. I assume the nested logit structure informs consumers sub-segment decisions, and that within sub-segments consumers care only about the tradeoff between fuel efficiency and other quality. This tradeoff can be captured by the single choice variable for firms of *mpg*. Now with the firms' characteristic choice reduced to one element, there is a clear, exogenous shifter (gas price) of this choice variable.

There are two costs of this approach. First, because I limit firms from introducing new automobiles, my model's predictions are only sensible in the medium run. Second, there is some loss of richness in the demand system by not retaining 6 preference characteristics. The benefit of this restriction, though, is that I have a clear instrument for firms' product. Also, I have a fully closed, endogenous system, rather than a mix of exogenous and endogenous portions of choice which lead to unrealistic parameter estimates. As well, the two costs of the approach may not be too large. In the first place, a larger, more complicated game admittedly happens at the five to ten year horizon in the auto industry that my model does not attempt to capture. The process of planning, investment, introductions, product launch, etc. is important but separate from this analysis. Second, the nested logit is indeed a useful demand system to apply to the automotive industry. A nested logit works well when the nesting groups are relevant to actual consumers' decisions, and the auto trade press as well as previous research have indicated that this is, at least to some extent, the case with autos.²³

²²This requires more significant design changes than simply modifying fuel efficiency within a sub-segment.

²³Goldberg (1995) and Goldberg (1998).

Firms face constant marginal costs of producing automobile j :

$$\ln(mc_j) = mc(mpg_j, X_{jt}, \omega_j, t) \quad (7)$$

X is the same as in demand: a set of exogenous controls for automotive sub-segment, segment interacted with fuel efficiency, and various macroeconomic variables. ω_j is the unobserved product cost (unobserved but inferred from the data). The specific functional form I use in the estimation results is:

$$\begin{aligned} \ln(mc_j) = & \gamma_1 + \gamma_2(t - t_0) + \gamma_3mpg_j + \gamma_4mpg_j(t - t_0) \\ & + \gamma_5(mpg - \overline{mpg}_{ss})^2 + \Gamma X_{jt} + \omega_j \end{aligned} \quad (8)$$

Because firms locate along an isocost curve, the standard interpretations of cost parameters do not apply. For instance, I would not necessarily expect $\gamma_3 > 0$ as would normally occur in an empirical model. The reason is that *within* a sub-segment, cars with relatively more fuel efficiency may *not* be more expensive to produce, precisely because they have less quality. The parameters reflect the shape of the isocost curve. Time trends are added to allow for technological improvement, which may change the curvature of the isocost curve over time. \overline{mpg}_{ss} is the average fuel efficiency in a sub-segment, and γ_5 is meant to capture that cars at extremes (either of high efficiency or low efficiency) may have different costs than more “average” cars in the sub-segment.

Note that ω_j are used in estimation moments, although, again, less restrictively than in previous work.

3.3 Modeling CAFE Standards

The federal government’s fines for violating CAFE standards in a given year are \$50 per car sold per 1 mpg below the standard that the firm average falls. It is widely thought, however, that there may be additional reputational and political

costs to American manufacturers of violating CAFE standards. The Big Three²⁴ have influence over the CAFE standards through their relationship with the federal government. The thought is that the manufacturers may stand to risk this influence by violating the standards, or might lose reputation with American consumers by violating our own, domestic energy policy. I leave the fine (λ in equations (5) and (6)) as a shadow cost to be estimated. The estimation results thus speak to whether the effective costs are higher than the costs of the fines alone.

I assume CAFE binds on all American manufacturers, in all years after 1977, and on no other manufacturers in any years. This is not exactly true, but it is a reasonable approximation of an elaborate system of carry-overs and credits used in calculating CAFE fines. It may be exactly true for Asian manufacturers. They have never paid fines, and have never been particularly close to the standards. Some European firms have paid \$50 per car-*mpg* fines. (In fact, European firms are the only firms to have ever paid CAFE fines.) However, for these European firms there may be less political pressure to abide by U.S. domestic energy policy than there is for U.S. firms. These fines to European manufacturers have tended to be on small fleets of luxury and sports cars where the pecuniary penalties are a small portion of the list price. American firms, in contrast to both Asian and European manufacturers, are likely to have been more constrained over the years. Not only have they hovered close to the regulation, they may also have faced unique pressure to abide by the standards because they are domestic firms. No American manufacturer has ever actually paid a fine, but this is due to the system of carry-overs and credits from year to year. The standard is likely to have essentially been binding (or near binding) for the whole time period. For simplicity, I simply assume a shadow cost on American manufacturers after 1977, and allow the parameter estimation to speak to the strength of the constraint.

4 Data & Industry

The model is estimated using data on all new automobiles sold in the US, and macroeconomic data, from 1971-2007. I have data on all sales and characteristics of passenger vehicles during that time period with two exceptions. First, I'm missing data from

²⁴General Motors, Ford, and Chrysler are the three American manufacturers.

1991-1995. The collection of the earlier year data stopped in 1990, and the electronic data-keeping by Wards Automotive did not begin until 1996. Second, I'm missing truck data for the early years (1971-1990), which were not collected with the original data set.²⁵

The early year data, 1971-1990, are the same data used in Berry, Levinsohn, and Pakes (1995), and were generously shared with me. These data are from Automotive News Market Data Book, and contain information on sales and base model characteristics of all passenger cars sold in the U.S. I supplement these with segment information that I collected from Wards Automotive Yearbooks. The later year data, 1996-2007, are from Wards Automotive. They contain sales, characteristics, and segment data on all passenger vehicles sold in the U.S. during those years. Table 4 shows popular cars in the various sub-segments for the later years.

Prices are list prices, fuel efficiency is city fuel efficiency as measured by the EPA (U.S. Environmental Protection Agency). The unit of observation is the vehicle model, the level at which sales data are reported. I use the characteristics of the base model for each model-year.²⁶

I also collect macroeconomic data from various government agencies: the Energy Information Administration (gas prices), Bureau of Labor Statistics (CPI, number of households, unemployment, income distribution), and the Bureau of Economic Analysis (GDP and GDP growth).

Table 3 contains summary statistics of both data sources. Sample average *mpg* is 20.7, higher for Cars and lower for Trucks and Utility Vehicles. There is substantial variation in the gas price, both in levels and in changes. Note that in 2008 (outside the sample summarized in the Table) the gas price jumped from \$2.80 to \$3.43 (through August), which is above the highest real gas price of the sample. This jump also nearly equalled the highest change of the sample, which is \$0.68 in 1980. There are 4,820 model-years over the course of 32 years, for an average of 151 models offered each year. There are 3 vehicle types, 9 vehicle segments, and 28 vehicle sub-segments.

²⁵I intend to collect the missing years of data to supplement the analysis.

²⁶Models have even finer distinctions of various trim levels. These offer slightly different configurations of engines and accessories. I ignore this in my analysis, as I do not have sales data at this level. The base model is usually the most fuel efficient trim of each model.

Figure 5 shows trends in the industry over time. The number of U.S. households (M), the market size, is the bottom line on the chart. It has grown at a relatively constant rate throughout the sample. Aggregate sales (middle line on the chart) has grown at more or less the same rate, albeit with substantially more variation around the trend. However, the number of vehicles (top line of the chart) has grown at a faster rate. The number of model offerings and even segments have proliferated throughout the sample period.²⁷

Figure 6 shows the industry's response to the gas price and regulation. The gas price is the solid line which moves by itself. The two dotted lines are the CAFE standards for passenger cars and light trucks (the standards apply separately for these groups). The solid lines which track them are the respective sales-weighted fuel efficiencies for the vehicle types. There have been 3 major gas price increases in the sample: 1973 (the OPEC embargo), 1979 (the Iranian Revolution), and 2007-2008 (supply interruptions and growing global demand). There have also been periods of historically low fuel prices in the late 1990s and early 2000s. This amounts to large variation in the gas price over the sample. CAFE standards were enacted in 1975 in response to the 1973 energy crisis. They imposed a standard which rose gradually until the early 1980s, and has remained largely unrevised until 2008. Recently the federal government has stipulated higher standards by 2020, and a stricter treatment of SUVs and trucks in calculating averages.²⁸ Note that average fuel efficiency has tended to outpace regulation during the gas price highs (1979 and into the 2000s), but the regulations appear to have mattered during times of low gas prices in the mid 1980s and 1990s. This is especially true for American manufacturers, although that is not directly observable from this aggregate chart.

Figure 7 shows yearly characteristics for one long-lived base model in the sample, the Toyota Celica. Note the pattern of inverse movements in *mpg* and *hp* or *weight*, as well as an *mpg* response to the gas price. However, these patterns do not necessarily hold exactly with each movement. Note also that product characteristics change quite frequently, so it is reasonable to think that firms are able to change these characteristics yearly by making adjustments to engines, materials, amenities, etc.

²⁷The entire SUV and CUV segments did not exist before the mid 1990s, but in 2007 comprised 30% of sales.

²⁸All firms' vehicle averages, both for passenger cars and for light trucks, must exceed 35 *mpg* by 2020. As well, vehicles over 8,500 lbs. of gross vehicle weight are no longer exempt from the standards. The new exemption is 10,000 lbs.

My utility specification (equation (2)) effectively assumes that consumers expectations of the future gas price are summarized by the current price of gas. The reason I assume this is that gas prices are notoriously difficult to predict. Oil futures are a relatively recent instrument, they do not exist for time periods longer than 12 months, and are usually quite close to the spot price. Another way of saying this is that spot prices incorporate almost all information about future oil prices. Consumers may still, of course, react to oil prices in other ways besides considering current levels, but Busse, Knittel, and Zettelmeyer (2008) show that sales-weighted fuel efficiency purchases track quite closely with the current gas price, not outpacing it nor lagging it. So this is the specification I use in the estimation results, and I take up other specifications in the robustness section (7).

5 Estimation

The estimation routine is Generalized Method of Moments, and the moments are constructed from the timing of the model.

5.1 Estimation Moments I

To help understand the estimation moments, it is useful to refer again to Figure 4, the model timing. The two information sets defined in the Figure are snapshots of information at various points in the year leading up to production. Loosely speaking, these are:

$$H_{t-1} = \{\text{Information known about time } t \text{ at the "beginning" of time } t - 1\} \quad (9)$$

$$I_{t-1} = \{\text{Information known about time } t \text{ at the "end" of time } t - 1\} \quad (10)$$

Pick the example year $t=2008$. Let H_{t-1} be the knowledge at the beginning of 2007

about what the auto industry will look like in 2008. This includes the set of all 2008 models, their manufacturers, their sub-segments, and the 2007 gas price. H_{t-1} does *not* include the cost and demand shocks, *mpg* choices, or gas prices and macroeconomic variables that will prevail in the 2008 marketplace.

Then, still in 2007, the cost and demand shocks, ω and ξ , that will affect the 2008 market are revealed to all manufacturers on all models. They are common knowledge. In response to these, and still in 2007, *mpg* is chosen for all 2008 models. Now take another snapshot and redefine this information set as I_{t-1} . It contains H_{t-1} , but is also updated to include shocks and *mpg*. It still, however, does not include the final pieces of the 2008 market - gas prices and macroeconomic variables. So, more specifically:

$$H_{t-1} = \{1, p_{gas,t-1}, \text{ownership/sub-segment of year } t \text{ models}\} \quad (11)$$

$$I_{t-1} = \{1, p_{gas,t-1}, \text{ownership/sub-segment of year } t \text{ models}, \omega_t, \xi_t, mpg_t\} \quad (12)$$

ω_t , ξ_t , and mpg_t are vectors for the entire industry, observed by all participants. I include “1” in these information sets so that shocks that are uncorrelated with the sets will be mean zero.

Now, two implications of the timing of the model are that demand and cost shocks that will prevail upon the market in 2008 (ξ_t and ω_t) are uncorrelated with the exogenous and unchosen portions of the 2008 market (H_{t-1}):

$$E[\xi_{jt} * H_{t-1}] = 0 \quad (13)$$

$$E[\omega_{jt} * H_{t-1}] = 0 \quad (14)$$

Errors in this demand and supply structure should not be correlated with the structure itself, otherwise different parameters fit the model better or the model has not

controlled for some exogenous but relevant component of demand and cost. For example, if ξ is correlated with a certain sub-segment, then it is likely that the model has mis-estimated the sub-segment intercept in utility. These implications (equations (13)(14)) hold element by element for these two vectors - no vehicles' shocks should be correlated with the exogenous market structure.

Note that the *logic* of this assumption (shocks orthogonal to exogenous market structure) is similar to previous estimation assumptions for demand models. However, the formulation here is less restrictive, because I exclude endogenous choice variables from the instrument set. This means that in my model firms are able to observe market shocks when choosing their characteristics. The standard demand identifying assumption would be to add mpg_t into what I call H_{t-1} , which would mean firms are unaware of the cost and demands shocks when they commit to characteristic decisions. This scenario is implausible in the auto industry, both because ξ and ω are partially chosen by manufacturers, and also because automobile models survive multiple years and there is learning. For example, the Toyota Camry sells quite well relative to other cars with comparable characteristics. This indicates a high ξ for the Camry, and all auto firms are likely to be aware of this when they choose their own product characteristics.

5.2 Estimation Moments II

The second set of estimation moments is also an implication of the timing of the model. They are derived directly from the first order conditions on optimal firm choice of p and mpg from equations (5) and (6). Because the first order condition for mpg has an expectation, I have two choices. One is that I can simulate the empirical distribution of gas price change to approximate the integral in the expectation. Because this is computationally burdensome, however, I instead use a second method outlined in Hansen and Singleton (1982).

To understand the Hansen and Singleton moments, define:

$$R_{jt} = \frac{\partial \Pi_{ft}}{\partial mpg_{jt}} \quad (15)$$

$$R_{jt} = \sum_{r \in \mathfrak{S}_f} \left[s_j \left(\frac{\partial p_r}{\partial m_j} - \frac{\partial mc_r}{\partial m_j} \right) + (p_r - mc_r) \frac{\partial s_r}{\partial m_j} + \lambda \{bind\} \frac{\partial \overline{mpg}_f}{\partial m_j} \right] \quad (16)$$

R_{jt} is the derivative of the profit function with respect to mpg_{jt} . In Figure 4, this appears at the extreme right of the timetable. It is the object which is set to zero *in expectation* in the firms' first order condition on optimal mpg choice (equation (6)). But because the gas price changes after mpg_{jt} is chosen, R_{jt} will not necessarily equal zero in the year of sale. There will be ex post regret (thus the choice of the letter R) in the mpg_{jt} choice. Another way of saying this is that if firms could readjust mpg_{jt} during the year of sale, they would. This is *not* the same as saying that firms fail to optimize. Firms do optimize, they simply must do so at a time when not all the information they would wish to see (the updated gas price) has been revealed.

Now, with this definition, Hansen and Singleton (1982) note that an implication of any such stochastic optimization model is that:

$$E[R_{jt} * I_{t-1}] = 0 \quad (17)$$

This says there are no patterns of ex post regret (R_{jt}) which are systematic or in any way predictable at time $t - 1$. If there were such patterns, then this *would* imply that firms are not optimizing correctly. That would be inconsistent with the structure of the model, in which firms use as much information as they can in choosing optimal characteristics.

The benefit of this estimation moment is that it is far less computationally burdensome than simulating the empirical distribution of gas price changes. As mentioned, I do experiment with placing I_{t-3} and I_{t-5} into equation (17) to reflect the possibility that firms have to choose product characteristics more than 1 year before the year of sale. Note that all that is different in these two information sets is that they contain older gas prices ($p_{gas,t-3}$ and $p_{gas,t-5}$, as opposed to $p_{gas,t-1}$). I discuss the results of

this robustness test further in section 7.

5.3 Estimation Mechanics

The three moments, (13), (14) and (17) are stacked for estimation. The fourth estimation moment, the pricing first order condition in equation (5), is able to hold exactly for any value for any parameter values by $\xi(\theta)$ and $\omega(\theta)$ (Berry, 1994). So it does not enter the minimization function explicitly, although implicitly it affects the other moments. Estimation is by 2-stage GMM (Generalized Method of Moments). Parameter values are chosen to minimize the following objective function:

$$\min_{\theta} \begin{pmatrix} \xi_t(\theta) * H_{t-1} \\ \omega_t(\theta) * H_{t-1} \\ R_{jt}(\theta) * I_{t-1}(\theta) \end{pmatrix} W_n \begin{pmatrix} \xi_t(\theta) * H_{t-1} \\ \omega_t(\theta) * H_{t-1} \\ R_{jt}(\theta) * I_{t-1}(\theta) \end{pmatrix}' \quad (18)$$

I choose starting values for demand search parameters based on Two Stage Least Squares Estimation using just the first two moments. I then do a first stage of GMM where W_n is a block diagonal matrix containing the inverse of the instrument matrices H_{t-1} , H_{t-1} , and $I_{t-1}(\theta)$.²⁹ The weight matrix must be continuously updated because $I_{t-1}(\theta)$ is a function of θ by virtue of the fact that it contains $\xi(\theta)$ and $\omega(\theta)$. After this first stage, I do a second stage of GMM where W_n is the standard block diagonal matrix containing the inverses of the variance-covariance matrices of the moments themselves.

I do ignore two equilibrium effects in the calculation of R_{jt} , equation (16), because they are likely to be small in magnitude and yet add considerable computational complexity. The first is the $\frac{\partial \mathbf{p}}{\partial m p g_j}$ term - the effect that changing one's fuel efficiency will have on the equilibrium prices (of all vehicles) that are played in the subsequent game. I assume this term to be zero. The second omission is that in calculating $\frac{\partial m p g_j}{\partial m_j}$, I ignore $\frac{\partial \mathbf{q}}{\partial m p g_j}$ - the effect that changing one's fuel efficiency will have on equilibrium quantities in the downstream game.³⁰ In other words, I ignore effects of

²⁹Or, this can be thought of as an identity weight matrix and normalized instruments.

³⁰This term also includes another instance of the first ignored term, $\frac{\partial \mathbf{p}}{\partial m p g_j}$.

mpg choices on the equilibrium of the downstream game. This means that I am only approximating the subgame perfect Nash equilibrium, but this approximation is a reasonable price to pay for tractability. These terms are likely to be small because of the number of models in the game, and that would add considerable computational burden.

6 Estimation Results

The results of the GMM estimation are reported in Table 5. Overall they show a good model fit with sensible and statistically significant parameter estimates. The nesting parameters (σ s) are between one and zero and descending. They are quite close to one which indicates the nests are relevant to decisions and that there is not much substitution across car types. The three main preference parameters, α , β_d , and β_m , are statistically significant and negative as expected.³¹

Fuel economy preference is allowed to vary by segment. The omitted segment from these interaction terms is Large Vans. Note that all segments have a net preference for fuel economy, ie a negative net coefficient on dpm . Conditional on sub-segment of purchase, Utility Vehicle purchasers are actually the *most* sensitive to fuel economy, trucks and vans tend to be in the middle, and cars are actually the least sensitive to fuel economy. This may seem counterintuitive at first, but it is sensible *conditional* on sub-segment of purchase. A driver saves more on fuel costs (driving habits equal) by increasing *mpg* from 16 to 18 (SUVs) than from 26 to 28 (Middle Cars). Mathematically, this is $\frac{\partial^2 u}{\partial mpg \partial mpg} < 0$, that the marginal utility of *mpg* is decreasing in *mpg*. The massive decline in SUV purchases under the high gas prices of 2008 corroborate this result rather than refute it. Declines in car purchases have been smaller.

To interpret the magnitudes of these fuel economy parameter estimates, I have calculated willingness-to-pay (WTP) in the right panel of Table 5. These are the WTP for a hypothetical increase in fuel efficiency from 25*mpg* to 30*mpg*. I have calculated this for 3 different automotive segments, the lowest, median, and highest WTP for fuel

³¹Again, note that the way these variables are defined, they are all economic “bad”s. dpm is the negative of fuel economy, mpg is the negative of quality, and p always enters utility negatively.

economy (Luxury Cars, Compact Cars, and SUVs). I have also calculated this WTP under average and high gas prices (\$2.00 and \$3.50). WTP is higher during times of high gas price because the marginal benefit of fuel efficiency, $\frac{\partial u}{\partial mpg_j}$, is increasing in the price of gas. With \$2.00 gas, WTP for the increase is \$4,098 for Luxury Cars, \$7,377 for Compact Cars, and \$11,749 for SUVs. With \$3.50 gas, these numbers increase to \$7,172, \$12,910, and \$20,560.

Moving back to the parameter estimates in the left panel of Table 5, the time trend interacted with UV (Utility Vehicle) captures the growing popularity of SUVs and CUVs as they are introduced throughout the 1990s and 2000s. It is generally best to assume that preferences are stable, rather than having trends such as this. However, the dramatic growth in Utility Vehicle market share (from 0 to 30% in a decade) necessitates a trend because it happened to coincide with the steadily falling gas prices. Because measured characteristics of the Utility Vehicle segment did not change over time, the only way the model can rationalize this pattern (without a time trend) is to estimate that Utility Vehicle purchasers *dislike* fuel economy. As I have described, however, Utility Vehicle purchasers are actually the most sensitive to fuel economy given the 2008 evidence. So the “insensitive” estimate is not plausible, and is better represented with a time trend on UV preference. This restores the result that consumers value fuel economy, especially in the UV segment.

GDP variables are statistically significant, especially GDP growth per capita. Asian and European manufacturers provide slightly higher utility on average. “autonews” is a dummy indicating the variable is from the early years of the data when Truck/Van data are not collected. Cars are mechanically and artificially ascribed to have “higher” market shares, and this is captured by the positive coefficient. The positive interaction of “autonews” and fuel economy reflects that these are all car purchases, which are, on average, less sensitive to fuel economy.

The cost parameters, as I mentioned, do not have the standard “increasing” cost interpretation. The negative coefficient on *mpg*, for instance, does not indicate that it costs less to add fuel efficiency to a car, all else equal. Instead, it means that conditional on a sub-segment, cars that have relatively more fuel efficiency and relatively less other quality (higher and leftwards on the isocost curve in Figure 1) are less expensive to produce. Over time, however, these relatively more fuel efficient cars

are becoming relatively more expensive to produce (or, providing quality is becoming relatively less expensive). Segment dummies, perhaps surprisingly are not that significant so I left them out of the specification. There is overlap in prices across segments which mutes the importance of the fixed effect. Also, segment is correlated with *mpg*, so including the controls tended to make the *mpg* term less significant. I left the terms out in the final specification.

The parameter estimate on the CAFE standard indicates that the shadow cost to American manufacturers of violating the CAFE standards is \$348 per *mpg* per vehicle per year. This is significantly larger than the non-compliance fine of \$50. A common hypothesis is that these shadow costs are due to political and reputational pressure, either from the federal government and/or domestic consumers. In any case, the parameter estimates are consistent with the notion that American manufacturers do, indeed, feel more constrained by the regulations than just the pecuniary penalty. They also suggests that during this time period domestic manufacturers would have set lower fuel efficiency in the absence of the CAFE standards.

The model implies marginal costs and markups. The average absolute markup (Figure 8) is \$1,706, and the average Lerner index (Figure 9) is .075, or 7.5%. These markups are somewhat smaller than in, for example, Berry, Levinsohn, and Pakes (1995). This may be due to the lack of random coefficients on preference parameters, which tend to create pockets of less elastic consumers. In the future I intend to add random coefficients on elements of the utility function in equation (2). The minimum implied marginal cost is \$4,600.

Allowing for correlation between shocks and characteristics proves to have quantitative significance. The correlation between the demand or cost shocks and the other endogenous choice variables is reported in Table 1. Previous literature had corrected for the endogeneity of these shocks with price (second column), however, they had not corrected for the endogeneity with the fuel efficiency choice (first column). The correlations I allow in the first column are not as strong as the correlations already corrected for in the literature (the second column), but they are still sizeable. The sign of the first column implies that cars with high demand (and/or high cost) tend to have less fuel efficiency and more other quality.

Table 1:

Estimated Correlations

	mpg	ρ	ω
ξ	-0.28	0.45	0.25
ω	-0.26	0.61	

The GMM parameter estimates change little from GMM stage 1 to stage 2. Most parameters change between 0-10%.

7 Robustness

There are a number of robustness checks of the model that I run. First, one might assume that firms must set product characteristics for year t before year $t-1$. In other words, there may be a longer lead time in production. I have tested this alternative assumption by assuming product characteristics are set 3 and 5 years from the year of sale. This does not change the estimation results much. This may seem surprising as it is a large change (conceptually) to the supply side model. However, the only element of I_{t-1} that is different from I_{t-3} and I_{t-5} is the price of gas, and the price of gas is highly serially correlated. So the instrument set orthogonal to the ex post regret (equation (17)) is largely unchanged.

A second robustness check is whether the addition of the sub-segment level of the nest is too fine a partition for automobiles. When nesting groups are too finely partitioned, too much of the utility can end up being generated by just the within-nest shares. This can lead to minimal substitution across nests. I re-estimate the model omitting sub-segments, but the results also change little. The nesting parameters are still quite high. Even with 3-levels of nest there were on average 7 vehicles per sub-segment. Therefore I favor the 3-level nesting specification. As I will show in counterfactual section 8, there is differential substitution across groups, even with all 3 nesting levels.

A third robustness check is to incorporate consumers who are forward-looking with

respect to the price of gas. This is done with interactions of fuel efficiency preference and recent gas price movements. If upswings (for example) in recent gas prices make people more sensitive to fuel efficiency, then people are expecting the gas price to continue upwards, or else have some behavioral bias towards paying attention to gas price. If upswings in recent gas prices make people less sensitive to fuel efficiency, then they expect prices to mean revert. It turns out that in this sample there is a slight tendency to expect prices to mean revert, but it is not a large effect. I have not run the full model and predictions with this specification. Given the difficulty in predicting gas prices, I have elected to use the spot price model.

I have also experimented with many controls (in X) on both demand and cost. Many of these which were not significant, and I left them out of the final specification. These included other macroeconomic variables such as the Federal Funds Rate, unemployment, and election year dummies. The cost controls also included regions of manufacturer, regions of manufacturer interacted with fuel efficiency, sub-segment, and interactions of segment with *mpg*. None of them impacted the estimation results much, they were usually insignificant, and they added time to an already-high dimensional search, so I left them out of the final specification.

8 Analysis of the 2008 Gas Price Increase

In 2008 the gas price increased 23% through August. The model is estimated on data through 2007, and can therefore make 2008 out-of-sample predictions as to consumers' purchase responses. These numbers can be compared to 2008 year-to-date actual figures which are emerging month to month.

The results of the model's 2008 predictions are summarized in Table 6. These are predictions assuming that the final gas price increase over 2007 levels is 23%. I display the predictions at 3 levels of aggregation. The predictions are quite close to the market metrics emerging so far. Through August, actual sales were down 12% over 2007 levels. The model predicts this decline almost exactly, - at 11.9%.

The model also predicts the dramatic reductions in sales of fuel inefficient SUVs. As discussed in the results section (6), the model estimates indicate that purchasers of Utility Vehicles have the highest marginal utility of fuel efficiency *conditional on* the choice of the nest. This is likely why there have been such dramatic reductions in sales of SUV and truck models, some of which have declined over 60% from 2007 levels. I have not yet collected detailed data on 2008 sales, so my comparisons are against the emerging trade press, and are tentative. Compared to these reports, the model probably exaggerates when it predicts Large Luxury SUV sales to be down a full 90%. However, predicted reductions of Large and Middle Luxury SUVs at 66% and 78% seem to be in the correct ballpark.

One effect of these compositional sale changes is to dramatically increase sales-weighted fuel efficiency, as well. This happens both within types (especially UV's which are predicted to be up 45.3%) and in the aggregate (predicted to be up by 27.9%). As mentioned, I do not have exact year to date figures to which to compare these.

In subsequent autumn months (at the time of writing) gas prices are falling once again. This tends to stimulate auto sales. However, at the same time the macroeconomy is struggling which depresses sales. Gas prices and the macroeconomy are captured in the model, but this late data should provide a real test of the so I'm hopeful predictions will hold to year's end, but final figures are not in. If the predictions diverge, it will likely be due to the unique macroeconomy and credit situation after August 2008.

I also investigate vehicle offerings for 2009 and beyond. I demonstrate that firms have incentives to raise fuel efficiency in their future automobiles. Consider, for example, Figure 10 which analyzes the Toyota Avalon. I chart two lines. Both are derivatives of the Toyota profit function with respect to Avalon *mpg*.³² Both use 2007 market structure, but the left line is this derivative under 2007 gas prices while the right line is this derivative under 2008 gas prices. Note that both lines are decreasing - when simulated fuel efficiency is low, Toyota stands to gain profit by increasing fuel efficiency. Then, at some point, these derivatives cross zero - this is the point at which Toyota would have no ex post regret in their choice of *mpg_{Avalon}*. At fuel efficiencies

³²In fact, both of these are R_{jt} as in equation (15), although under different gas prices.

above this optimal, both profit functions decrease.³³

Comparing the two lines reveals the effect of the gas price increase of 2008. To start, notice that model implies the 2007 optimum to be 21.3 *mpg*. This is quite close to Avalon's actual 2007 *mpg* of 22. Now, when the gas price increases in 2008, this shifts the entire $\frac{\partial \pi}{\partial mpg}$ curve. The new optimum is 23.5 *mpg*. This is an increase of 2.2 *mpg*, or 10.3% over 2007 optimum. The gas price increase of 23% induces a 10.3% increase in optimal Avalon fuel efficiency.

I calculate this optimal change both in levels and in percentages for all 2007 models. A histogram of these implied changes can be found in figures 11 and 12. Note that the changes are somewhat heterogenous - some are even negative. The average implied optimal increase in fuel efficiency is 4 *mpg*, or 21.2%.

These figures should be treated cautiously. They measure incentives *holding all other firms fuel efficiencies fixed*. Of course, as each firm would change to a new optimum, all other firms' fuel efficiencies would not be held fixed. So this in no way represents an equilibrium. The equilibrium percentage increase in fuel efficiency could be larger or smaller than 20%, depending on whether fuel efficiencies are strategic substitutes or complements.

I have not yet computed the new equilibrium but am in the process of doing so. The equilibrium is not guaranteed to exist, but recent empirical applications indicate that an equilibrium may nonetheless be found (e.g. Lustig (2008)). The equilibrium involves finding a fixed point of optimal fuel efficiencies of all models.

It is also worth noting that these implied optimal changes are essentially uncorrelated with 2007 sales levels - the correlation coefficient is .06. At face value this might suggest that equilibrium sales-weighted fuel efficiency would be close to the unweighted change in fuel efficiency offerings. But, again, a new sales-weighted efficiency figure would of course also depend on finding the proper equilibrium.

Just how quickly firms could implement any these optimal changes is an open ques-

³³Note that the lines in the chart, which are derivatives themselves, also have negative derivative. This indicates that the the objective function (profit) is concave in *mpg* choice.

tion. To take the year-to-year timing of the model literally, they would be in place in 2009. However, this is likely too quick to realize all these changes. Also, at the time of writing (autumn 2008) the gas price is falling, and this may have already muted some of these incentives to increase efficiency. But the model does indicate that the gas price increase affects firms' optimal choices by providing (at least unilateral) incentives to increase fuel efficiency.

9 Conclusion

The auto industry is large, it consumes a significant portion of the nation's energy needs and contributes significantly to carbon emissions. Understanding the impacts that various policies and markets (particularly the gasoline market) have on the auto industry is crucial to informing energy policy.

I develop a model of the automobile industry to analyze these effects. In the model, firms provide more or less fuel efficiency depending on the exogenously changing gas price. The model contributes to the literature in three ways. First, unlike previous work I allow firms to choose product characteristics of their new vehicles. This allows me to analyze changes in the characteristics (in particular fuel efficiency) themselves. Second, I relax identifying assumptions commonly used in the empirical demand estimation. These identifying assumptions are restrictive and implausible in the automobile market, and I relax them by forming moments based on the timing of product selection. Third, I provide parameter estimates that imply that consumers care about fuel efficiency. Previous work has often had difficulty doing so, due to a negative correlation between the economic and quality effects of fuel efficiency for which I control.

The model can be used to analyze a number of market and policy changes. It is best adapted to addressing counterfactual gas price changes, which also have a direct and obvious analogue to gas price changes. The model could potentially be used to analyze changes in CAFE standards, although perhaps less effectively. The CAFE standards influence the industry on a time scale that is longer than what my model considers.

I estimate the model on the U.S. automobile industry. Parameter estimates imply that American manufacturers perceive significant costs to violating CAFE standards. These shadow costs are estimated to be \$348 per car-*mpg*, which is larger than the \$50 fine.

I also analyze the 2008 gas price increase. The model makes out-of-sample predictions that match 2008 figures well. I then demonstrate firms' incentives to raise fuel efficiency on the order of 20% in subsequent vehicle offerings (for 2009 and beyond). This figure should be treated cautiously, however, because it holds all other firms fuel efficiencies fixed. It is not an equilibrium. Whether the equilibrium percentage change is larger or smaller would depend on whether fuel efficiencies are strategic substitutes or complements. I am currently pursuing this new equilibrium of fuel efficiency choices.

References

- BERRY, S. (1994): “Estimating Discrete-Choice Models of Product Differentiation,” *RAND Journal of Economics*, 25, 242–242.
- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63, 841–841.
- (1999): “Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy,” *American Economic Review*, 89(3), 400–30.
- BUSSE, M., C. KNITTEL, AND F. ZETTELMEYER (2008): “Pain at the Pump: How Gasoline Prices Affect Automobile Purchasing,” *Working Paper*.
- CARDELL, N. (1997): “Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity,” *Econometric Theory*, 13, 185–213.
- CRANDALL, R., AND J. GRAHAM (1989): “Effect of Fuel Economy Standards on Automobile Safety, The,” *Journal of Law and Economics*, 32, 97.
- CRAWFORD, G., AND M. SHUM (2007): “Monopoly Quality Degradation and Regulation in Cable Television,” *The Journal of Law and Economics*, 50(1), 181–219.
- EPA (2008): “Environmental Protection Agency: Inventory of US Greenhouse Gas Emissions and Sinks, 1990 - 2006,” .
- FEDERAL RESERVE BOARD, C. (2003): “Predicting Oil Prices,” .
- GOLDBERG, P. (1995): “Product Differentiation and Oligopoly in International Markets: The Case of the US Automobile Industry,” *Econometrica*, 63, 891–891.
- (1998): “The Effects of the Corporate Average Fuel Efficiency Standards in the US,” *Journal of Industrial Economics*, 46(1), 1–33.
- GOLDBERG, P., AND F. VERBOVEN (2001): “The Evolution of Price Dispersion in the European Car Market,” *The Review of Economic Studies*, 68(4), 811–848.
- GRUENSPECHT, H. (1982): “Differentiated Regulation: The Case of Auto Emissions Standards,” *American Economic Review*, 72(2), 328–331.

- HANSEN, L., AND K. SINGLETON (1982): “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models,” *Econometrica*, 50(5), 1269–1286.
- HOTELLING, H. (1929): “Stability in Competition,” *Economic Journal*, 39(153), 41–57.
- KLEIT, A. (1990): “The effect of annual changes in automobile fuel economy standards,” *Journal of Regulatory Economics*, 2(2), 151–172.
- (2004): “Impacts of Long-Range Increases in the Fuel Economy (CAFE) Standard,” *Economic Inquiry*, 42(2), 279–294.
- LUSTIG, J. (2008): “The Welfare Effects of Adverse Selection in Privatized Medicare,” *Working Paper*.
- MAZZEO, M. (2002): “Product Choice and Oligopoly Market Structure,” *RAND Journal of Economics*, 33(2), 221–242.
- NEED (2008): “National Energy Education Development Project: Intermediate Energy InfoBook,” .
- PAKES, A., S. BERRY, AND J. LEVINSOHN (1993): “Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of Environmental Change,” *American Economic Review*, 83(2), 240–246.
- SEIM, K. (2006): “An empirical model of firm entry with endogenous product-type choices,” *RAND Journal of Economics*, 37(3), 619–640.
- SWEETING, A. (2007): “Dynamic Product Repositioning in Differentiated Product Industries: The Case of Format Switching in the Commercial Radio Industry,” *Forthcoming Econometrica*.
- THOMADSEN, R. (2005): “The Effect of Ownership Structure in Geographically Different industries,” *RAND Journal of Economics*, 36(4), 908–929.
- WU, T., AND A. MCCALLUM (2005): “Do Oil Futures Prices Help Predict Future Oil Prices?,” *Federal Reserve Board San Francisco Economic Letter*, (38).

Appendix A: Nested Logit Shares

This appendix details the nested logit errors, shares, and nesting parameters σ .

Rewrite the utility function in equation (2) as:

$$u_{ijt} = \delta_j + \tilde{\epsilon}_{ijt} \quad (19)$$

The nested logit error term $\tilde{\epsilon}_{ijt}$ is:

$$\tilde{\epsilon}_{ijt} = \epsilon_{i,v} + (1 - \sigma_v)\epsilon_{i,s} + (1 - \sigma_s)\epsilon_{i,ss} + (1 - \sigma_{ss})\epsilon_{ij} \quad (20)$$

Each individual i has a shock not only on the auto model (ϵ_{ij}), but also a shock common to all vehicles within the same sub-segment ($\epsilon_{i,ss}$), segment ($\epsilon_{i,s}$), and vehicle type ($\epsilon_{i,v}$). These error terms have a distribution such that the cumulative error term within any group follows the type-I extreme value distribution. Therefore, market shares for product j are:

$$s_j = \frac{e^{\frac{\delta_j}{1-\sigma_{ss}}}}{D_{ss}} * \frac{D_{ss}^{\frac{1-\sigma_{ss}}{1-\sigma_s}}}{D_s} * \frac{D_s^{\frac{1-\sigma_s}{1-\sigma_t}}}{D_t} * \frac{D_t^{1-\sigma_t}}{D_a} \quad (21)$$

where D 's are inclusive values for nests:

$$\begin{aligned} D_{ss} &= \sum_{j \in ss} e^{\frac{\delta_j}{1-\sigma_{ss}}} \\ D_s &= \sum_{ss \in s} D_{ss}^{\frac{1-\sigma_{ss}}{1-\sigma_s}} \\ D_t &= \sum_{s \in t} D_s^{\frac{1-\sigma_s}{1-\sigma_t}} \\ D_a &= \sum_t D_t^{1-\sigma_t} + 1 \end{aligned}$$

Taking logs and rearranging yields:

$$\ln \frac{s_j}{s_o} = \sigma_{ss} \ln \frac{s_j}{s_{ss}} + \sigma_s \ln \frac{s_{ss}}{s_s} + \sigma_v \ln \frac{s_s}{s_v} + \delta_j \quad (22)$$

δ_j customarily includes a demand error, ξ_j , that is used to form estimation moments.

Figure 2:

Nesting Structure

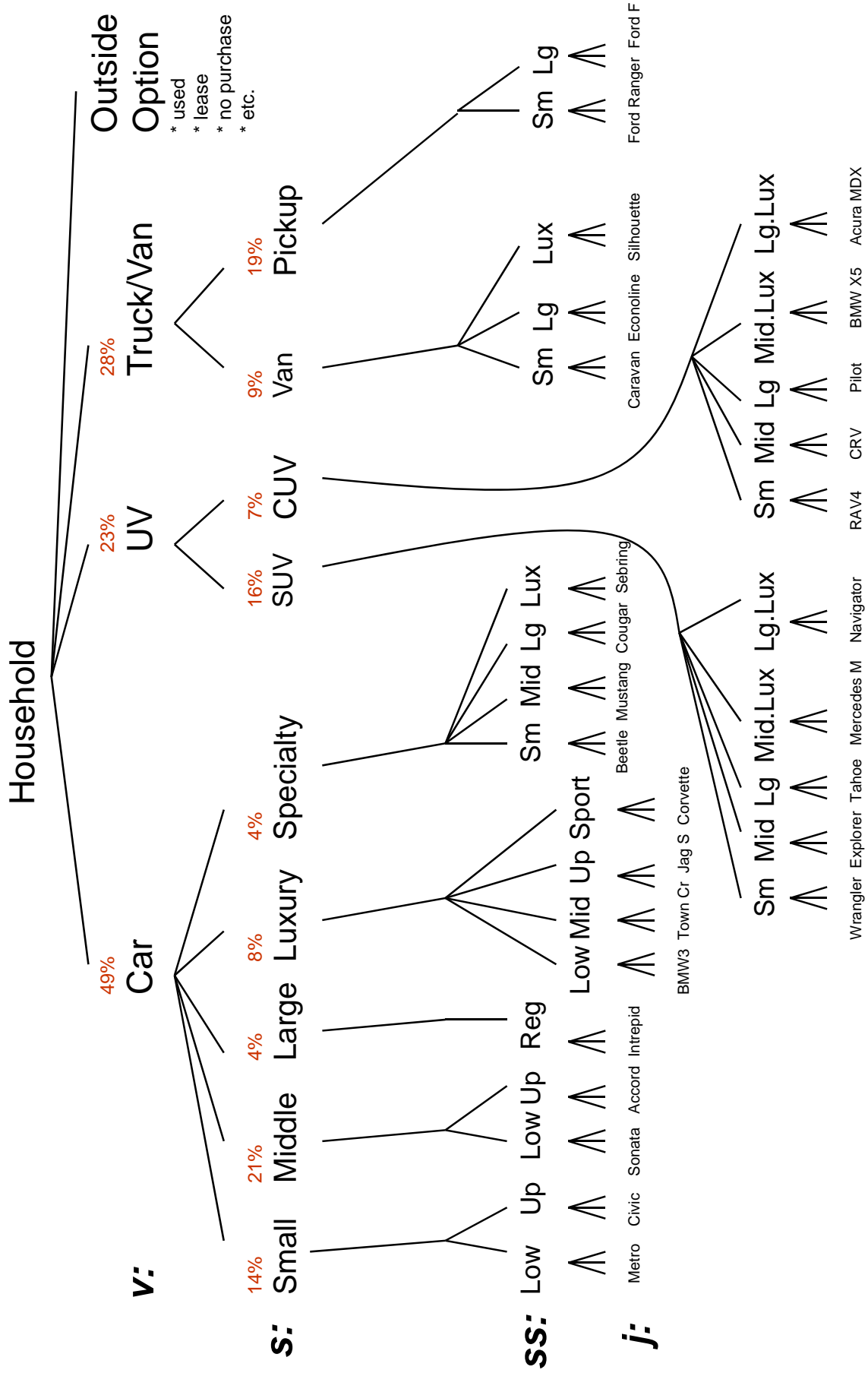


Figure 3:

**Technology Frontier
2007 U.S Vehicles**

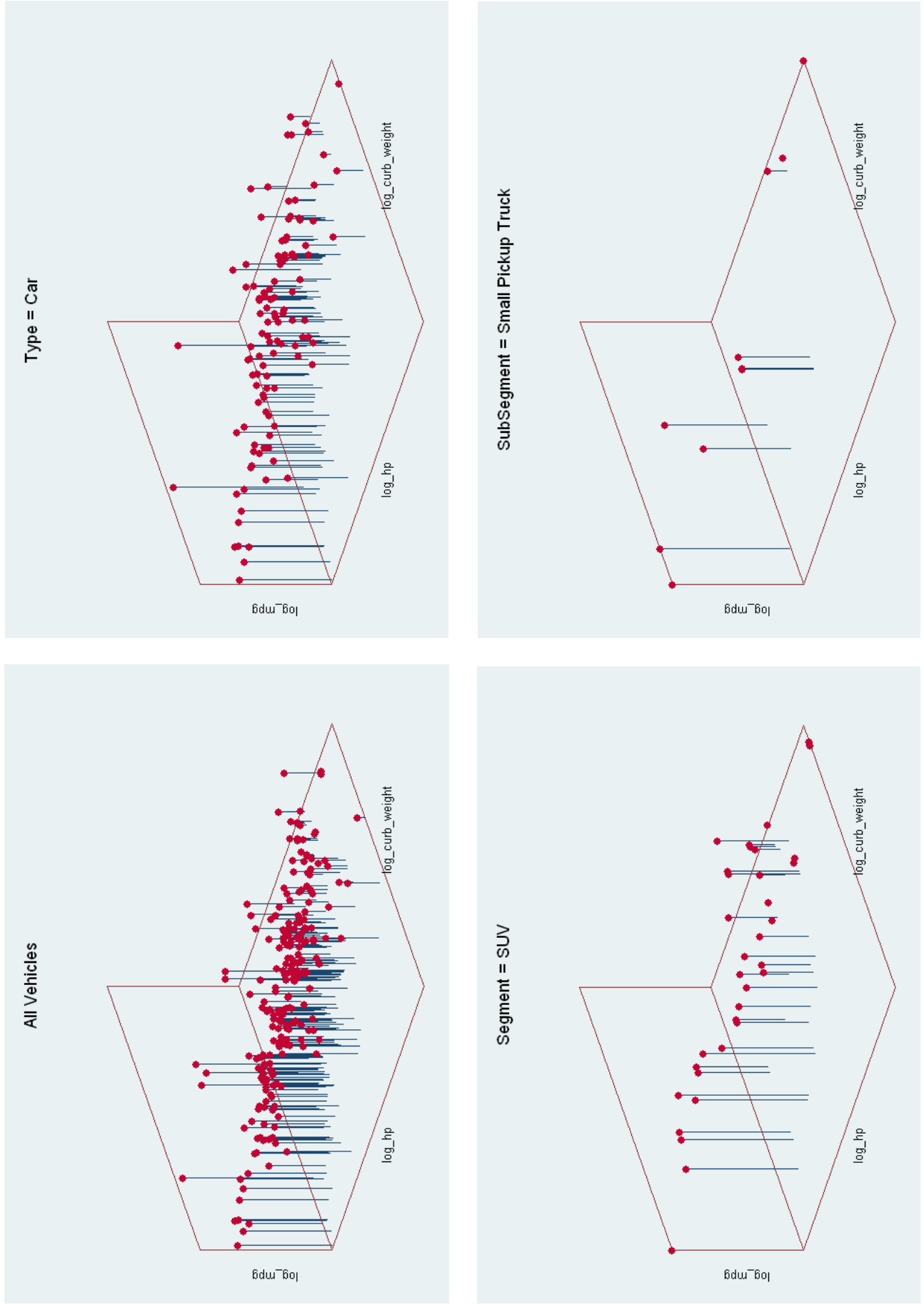


Figure 4:

Timing of Events

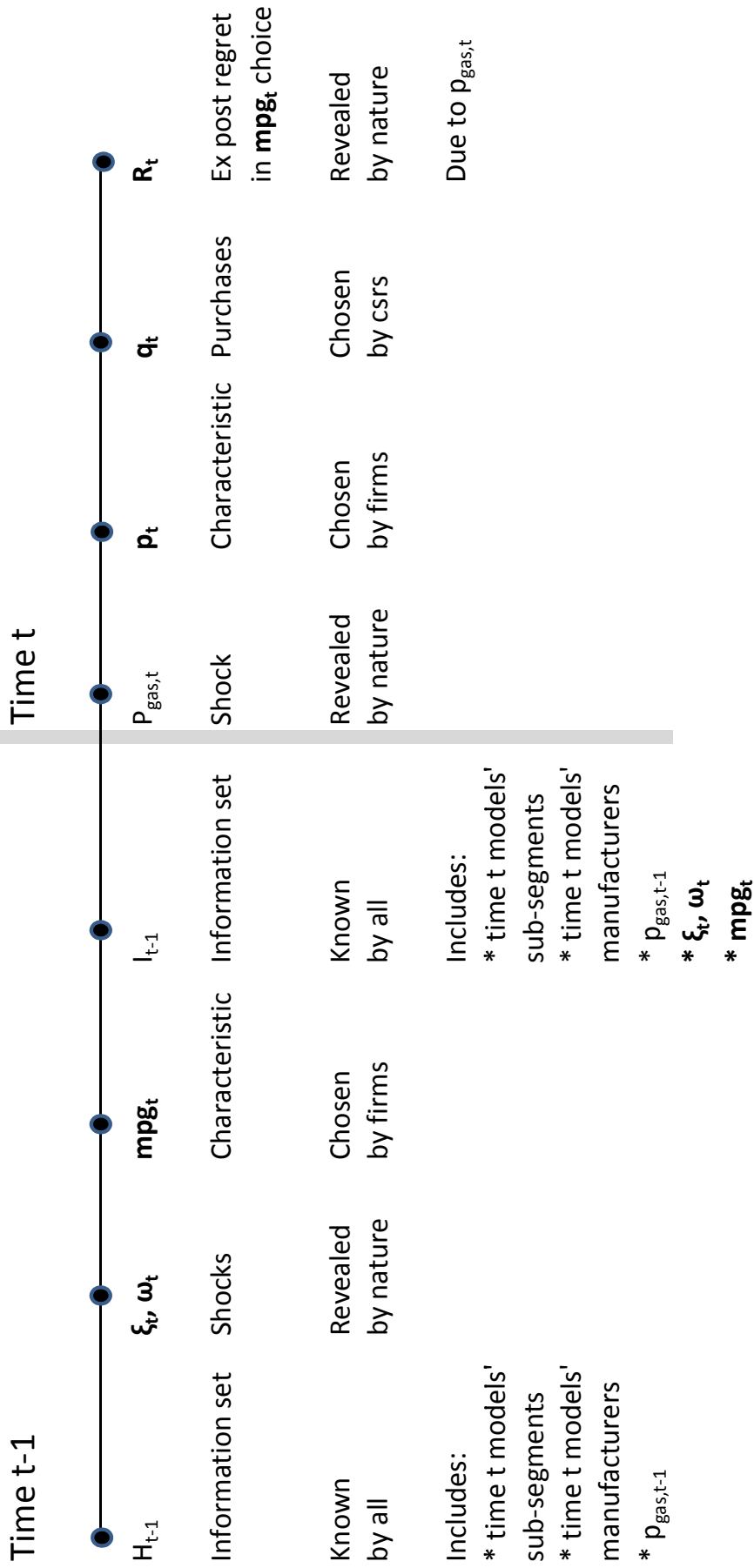
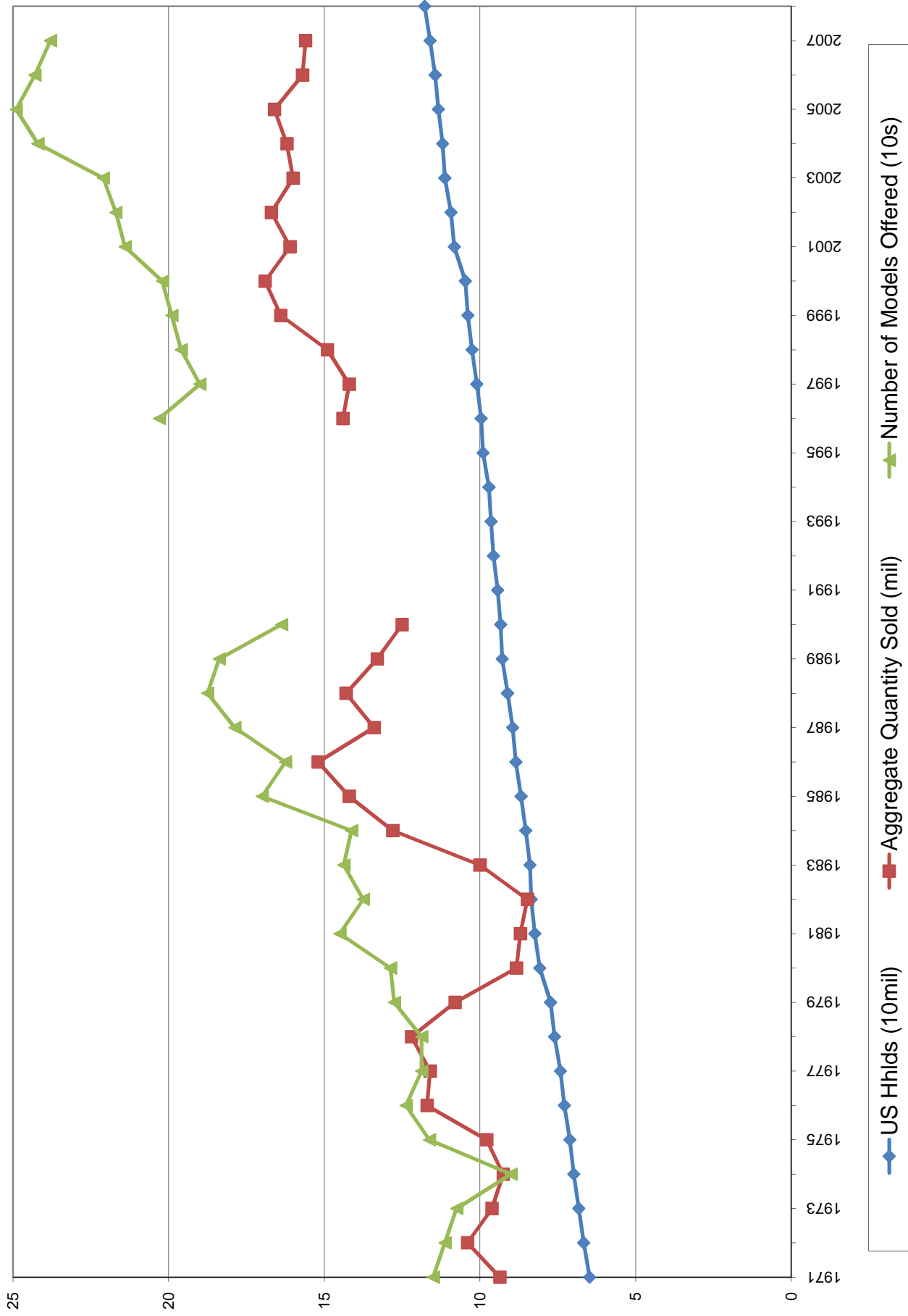


Figure 5:

Historical Industry Statistics

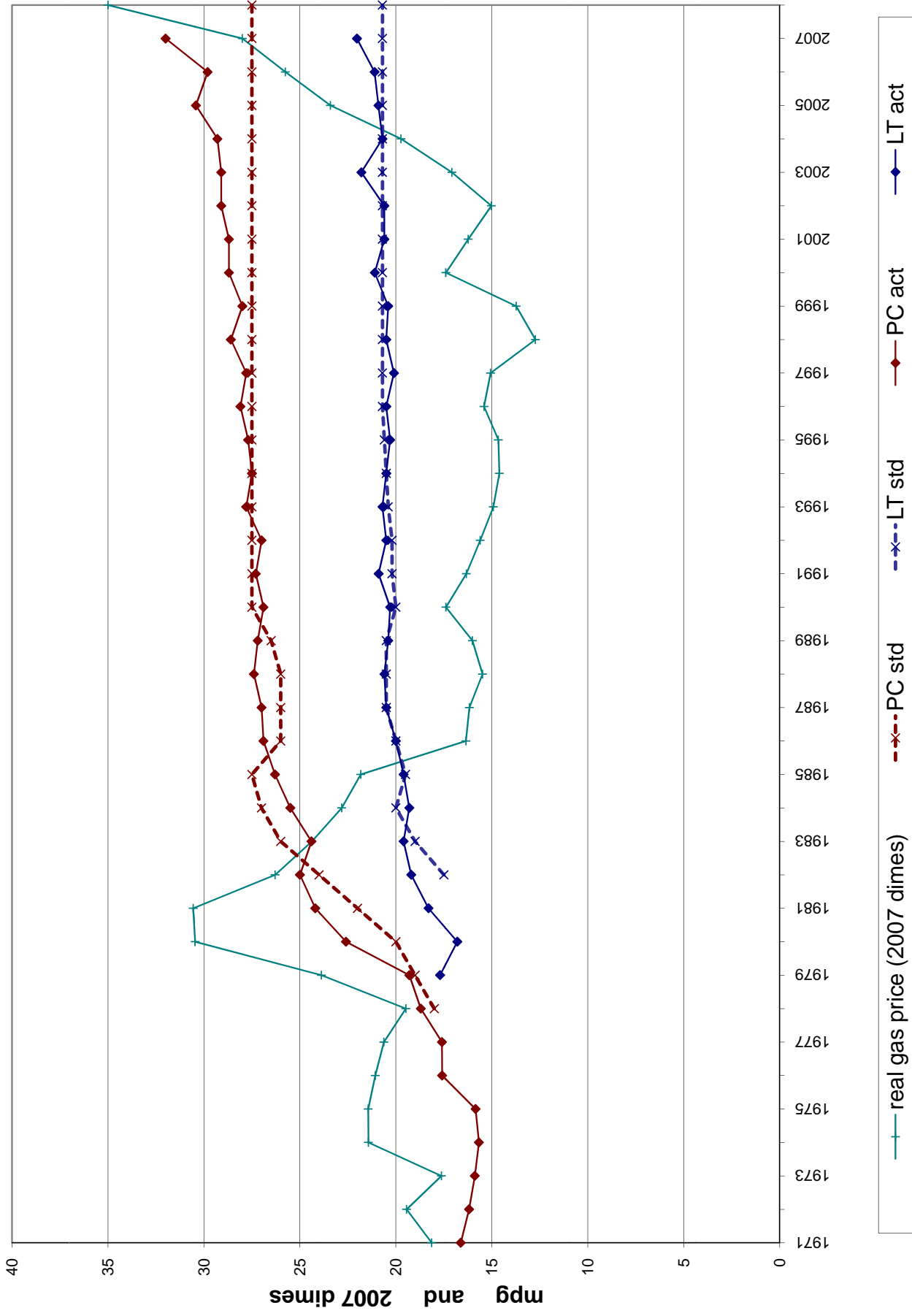


Pre-1991 Number of Models Offered is scaled up by sales to reflect missing truck models

Sources: BLS, Automotive News

Figure 6:

Historical Gas Price & Fuel Efficiency



PC = Passenger Cars, LT = Light Trucks

Sources: Wards, Automotive News, EIA, NHSTA

Figure 7:

Yearly Characteristic Changes Toyota Celica

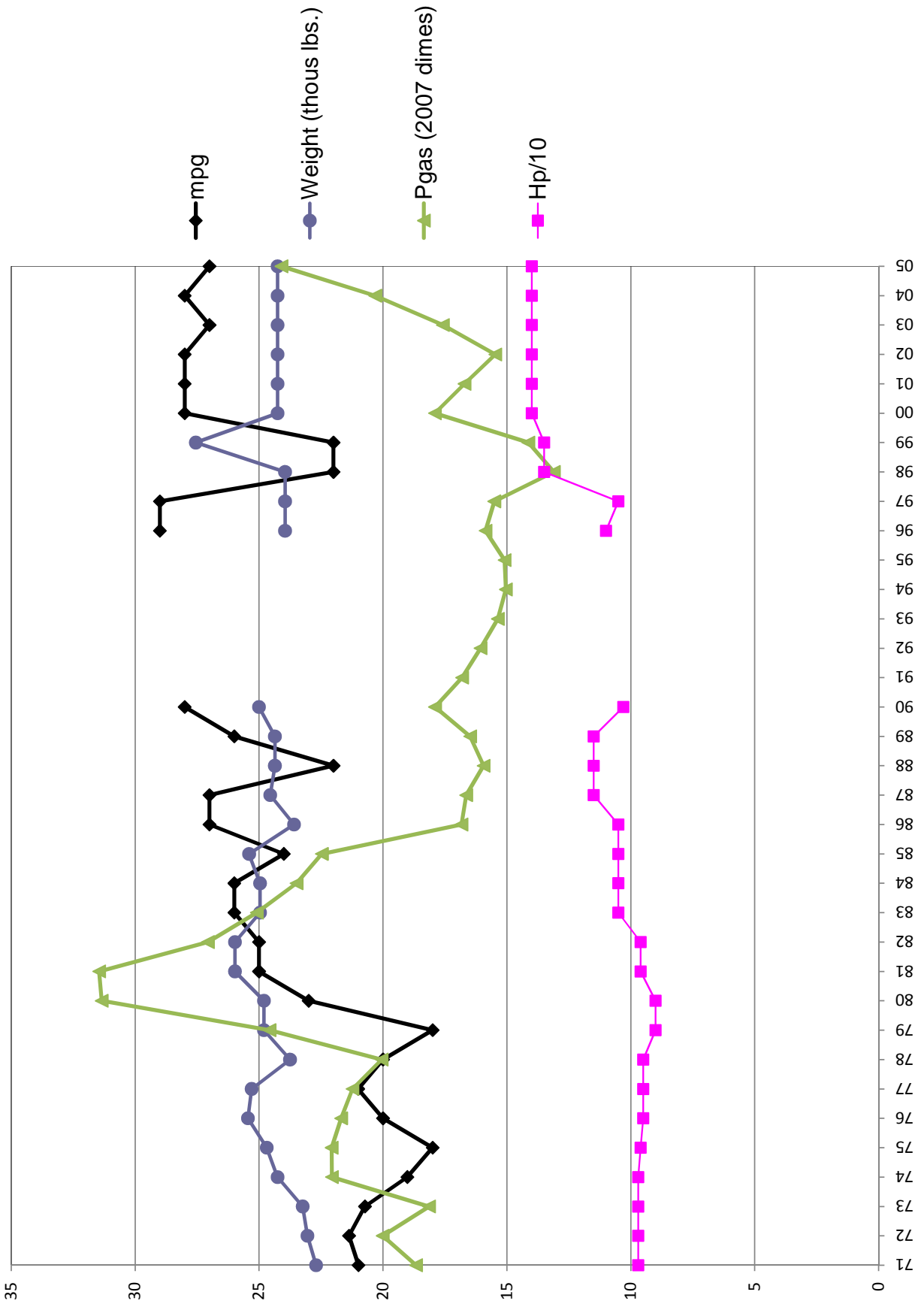
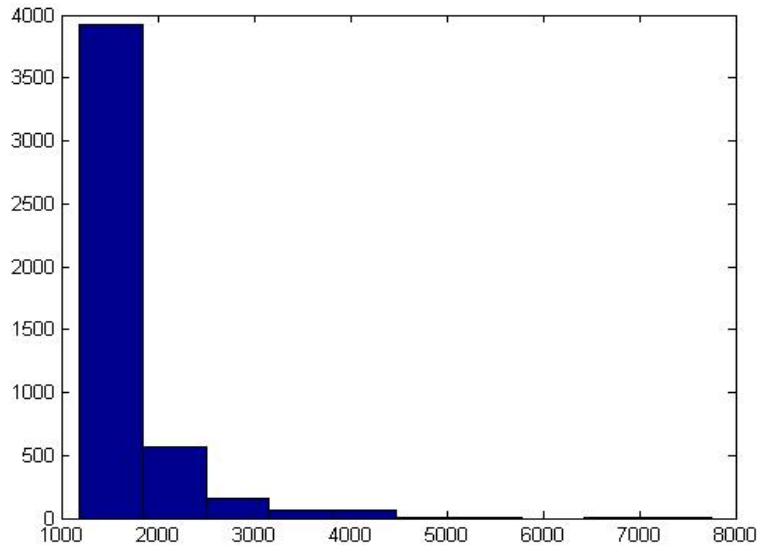


Figure 8:

Markup: $(p-mc)$
(Implied by Model)

Avg = +\$1,706



Note: Min. implied mc = \$4,600

Figure 9:

Markup %: $(p-mc)/p$
(Implied by Model)

Avg = +7.5%

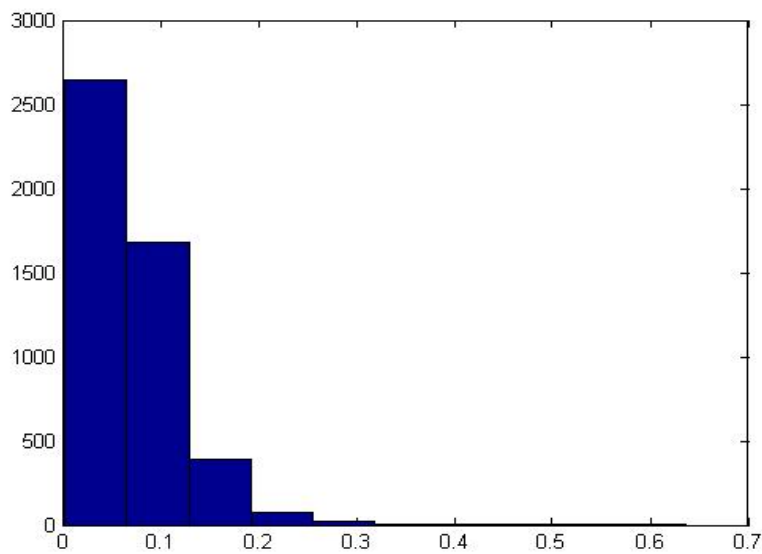


Figure 10:

Incentive to Change Fuel Efficiency Toyota Avalon

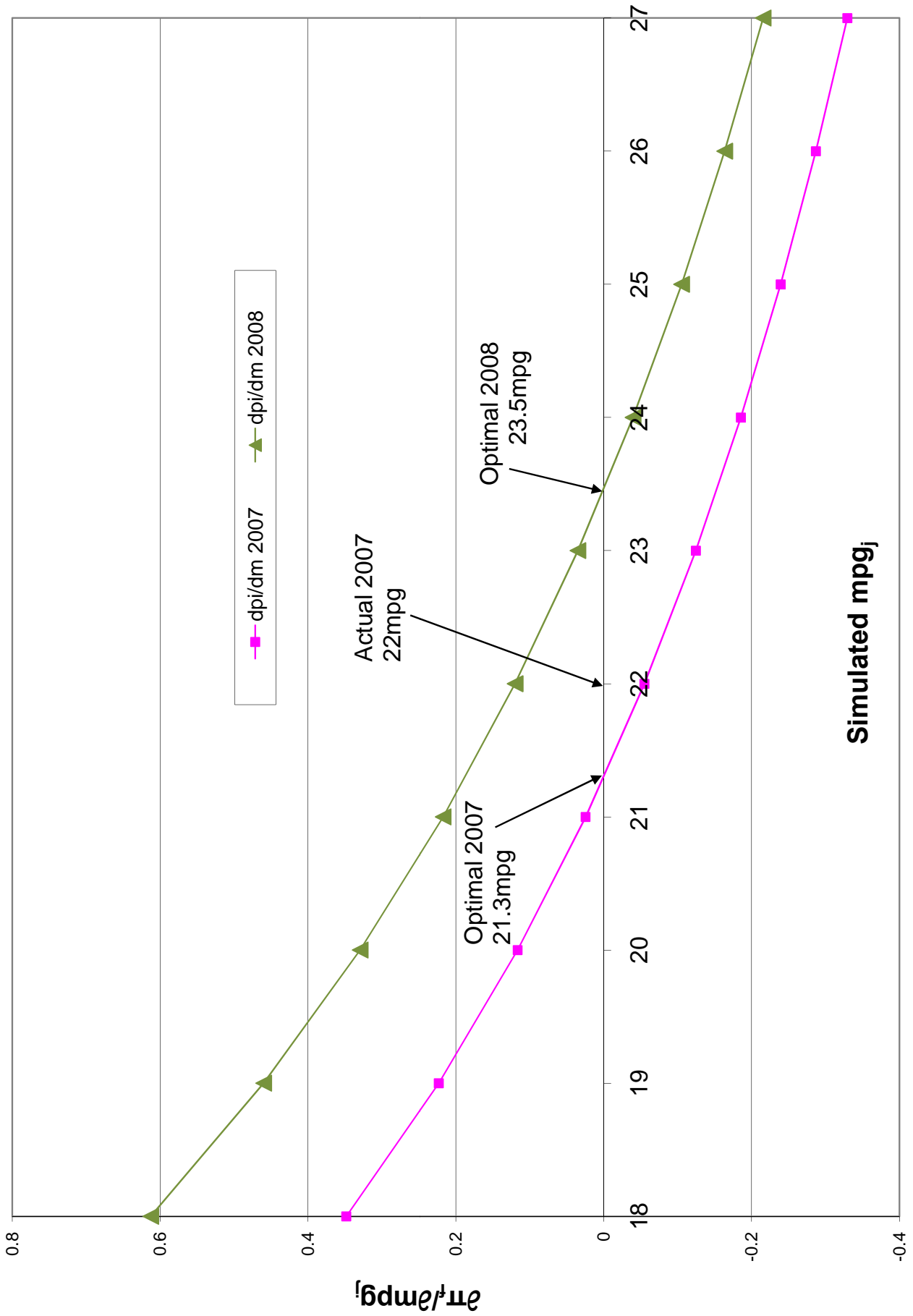


Figure 11:

Optimal Mpg Change
(Implied by Model)

Avg = +4.0mpg

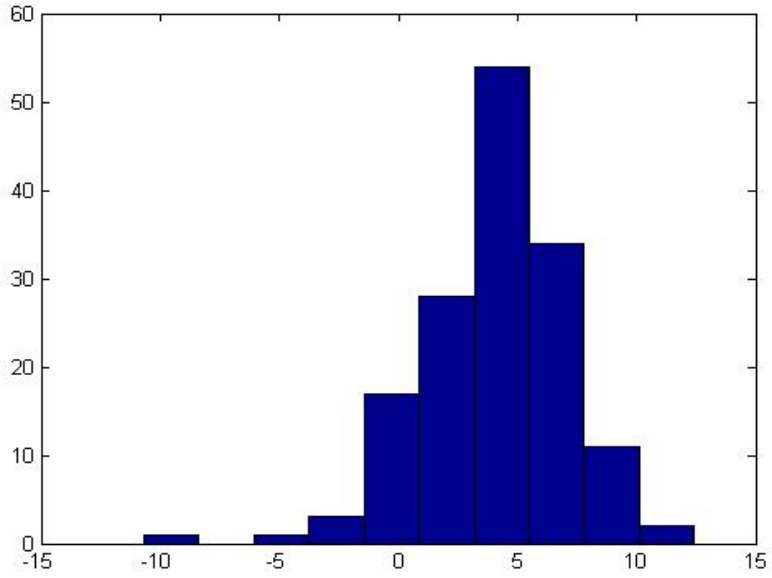


Figure 12:

Optimal Mpg Change (%)
(Implied by Model)

Avg = +21.2%

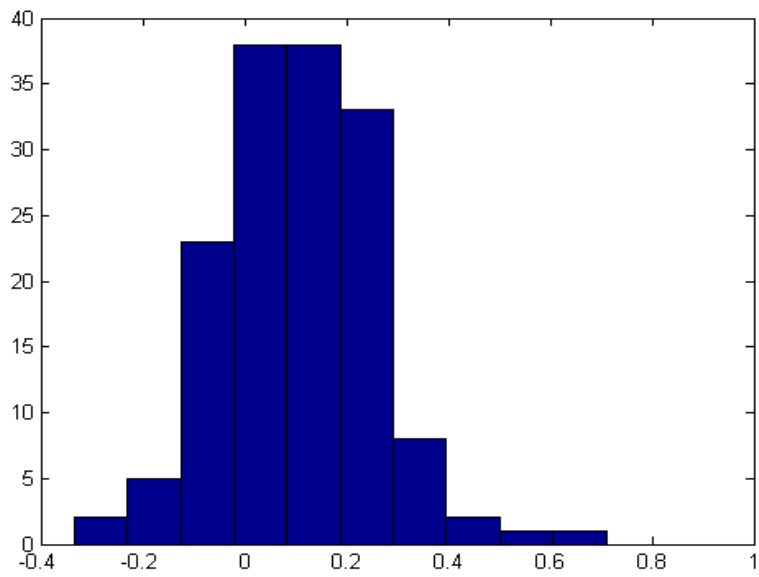


Table 2:

<u>OLS Regression</u>			
Dep Var = ln(mpg)	Coef	SE	t
ln(hp)	-0.279	0.008	-37
ln(curb_weight)	-0.596	0.012	-52
year	0.013	0.000	66
const	-16.682	0.397	-42
Nobs	4820		
R-sq	0.78		
R-sq between .4 and .85 for same regression within subsegment			

Table 3:

Summary Statistics				
Variable	Mean	Std. Dev.	Min	Max
Year	1992.5	11.1	1971	2007
Mpg	20.7	5.8	9.13	61
Dpm	\$0.10	\$0.04	\$0.03	\$0.24
Price	\$30,059	\$21,195	\$7,038	\$483,849
q	75,566	95,043	123	890,790
Price gas	\$2.02	\$0.48	\$1.31	\$3.15
Price gas change	\$0.05	\$0.24	-\$0.56	\$0.68
GDP Growth	3.1%	1.7%	-1.9%	7.2%
N Models in Year	150.6	55.9	72	249
N Model-Years	4,820			
N Years	32			
N Types	3			
N Segments	9			
N Sub-segments	28			
Note: Prices are Real 2007 Dollars				

Table 4:

Automotive Types, Segments and Sub-Segments

Segment Sub-Segment	% of		Popular Models	
	Model-Yrs	Sales		
Cars				
Small	10.9%	13.5%		
Lower Small	3.6%	1.1%	GMC Metro	Kia Sephia
Upper Small	7.3%	12.3%	Honda Civic	Ford Escort
Middle	11.5%	20.5%		
Lower Middle	4.7%	5.8%	Honda Civic	GMC Malibu
Upper Middle	6.8%	14.7%	Toyota Camry	Ford Taurus
Large	3.1%	3.8%		
Large Regular	3.1%	3.8%	GMC Lesabre	Chrysler Intrepid
Luxury	23.6%	7.6%		
Lower Luxury	7.6%	3.5%	BMW 3	Volvo 70
Middle Luxury	6.4%	2.6%	GMC Fleetwood Deville	GMC Park Avenue
Upper Luxury	3.8%	0.8%	Mercedes E	BMW 5
Luxury Sport	5.7%	0.7%	Corvette	Chrysler Crossfire
Specialty	10.8%	4.1%		
Small Specialty	3.3%	0.8%	Beetle II	Toyota Celica
Middle Specialty	4.5%	2.7%	Ford Mustang	GMC Camaro
Large Specialty	0.3%	0.1%	Ford Thunderbird	these are the only two cars in this class & they only exist 96-98
Luxury Specialty	2.6%	0.5%	Chrysler Sebring CNVT	Mercedes CLK
CUV	10.6%	6.5%		
Small CUV	1.4%	1.2%	Chrysler PT Cruiser	Toyota RAV4
Middle CUV	4.3%	3.5%	Ford Escape	GMC Equinox
Large CUV	1.0%	0.6%	Honda Pilot	Ford Freestyle
Middle Luxury CUV	3.0%	1.0%	Lexus RX330	Volvo XC90
Large Luxury CUV	0.9%	0.3%	Acura MDX	Mercedes R Class
SUV	13.0%	16.0%		
Small SUV	1.4%	0.9%	Chrysler Wrangler	GMC Tracker
Middle SUV	4.9%	9.3%	Ford Explorer	GMC Trailblazer
Large SUV	2.3%	4.5%	Ford Expedition	Chrysler Durango
Middle Luxury SUV	2.6%	0.7%	Mercedes M	GMC Bravada
Large Luxury SUV	1.9%	0.6%	Ford Navigator	GMC Hummer 2
Van	9.7%	9.4%		
Small Van	6.1%	6.6%	Chrysler Caravan	Chrysler Town & Country
Large Van	2.4%	2.2%	Ford Econoline	Toyota Sienna
Luxury Van	1.2%	0.6%	Chrysler Town & Country	Mazda MPV Pass. Van
Pickup	5.4%	18.5%		
Small Pickup	3.3%	5.3%	Ford Ranger	Toyota Tacoma
Large Pickup	2.1%	13.2%	Ford F Series	GMC Chevy CK Pickup
Other	1.4%	0.1%		
Comm. Chassis	1.4%	0.1%	Chevy P Model	Isuzu NPR
Utility Vehicles				
Trucks/Vans				

Table 5:

GMM Estimation Results

Implied WTP

Variable	coeff	se	t
p (\$10k)	-0.098	0.01	-7.2
$\ln(S_{j ss})$	0.97	0.01	87.1
$\ln(S_{ss s})$	0.94	0.01	72.1
$\ln(S_{s t})$	0.92	0.01	72.8
dpm	-4.16	0.49	-8.5
mpg	-0.029	0.00	-17.5
S_dums*dpm (9 params)	(-4.5, 1.2)		11.4*
$(t-t_0)*UV$	0.11	0.004	26.1
const	-2.87	0.13	-22.5
real GDP per cap grwth %	2.38	0.18	13.5
real GDP per cap (\$10k)	0.065	0.01	7.3
asian	0.011	0.01	1.2
euro	0.100	0.02	5.3
auto_news	0.254	0.03	9.1
auto_news*dpm	3.79	0.40	9.4
SS_dums (31 params)	(-3, 1)		9.8*
mpg	-0.075	0.01	-6.8
$(t-t_0)*mpg$	0.001	0.00	3.0
const	11.143	0.21	53.4
$(t-t_0)$	-0.0001	0.00	0.0
$(mpg-mpg_{ss})^2$	0.001	0.01	0.2
CAFE_bind	348	11.95	29.1
N		4,820	
R2 from 2sls on demand		0.95	

Implied Willingness-To-Pay for mpg increase
(from 25 mpg to 30 mpg)

pgas
\$2.00 \$3.50

Luxury Cars (Min WTP, $\beta_d = -3.0$) \$ 4,098 \$ 7,172
 Compact Cars (Med WTP, $\beta_d = -5.4$) \$ 7,377 \$ 12,910
 SUVs (Max WTP, $\beta_d = -8.6$) \$ 11,749 \$ 20,560

* Average of absolute value of t-stats

Model Predictions By 3 Aggregation Levels

Table 6:

2007 Figures - Actual
2008 Figures - Predicted

Type	Vehicle	Segment	Subsegment	# Models	Mpg Offered		Mpg Purchased (Sales Weighted)		Sales (1,000s)			
					Min	Max	2007	2008	Change (%)	2007	2008	Change (%)
All				237	13	60	22.6	28.9	27.9%	15,619	13,754	-11.9%
Car				119	14	60	26.8	33.5	24.8%	7,385	6,784	-8.1%
UV				85	13	36	20.0	29.0	45.3%	4,569	3,904	-14.6%
Truck/Van				33	14	24	17.4	18.8	7.8%	3,666	3,066	-16.4%
Car				6	27	34	31.3	32.3	3.0%	336	180	-46.3%
	Small	Lower		18	22	49	31.2	43.9	40.6%	1,878	1,223	-34.9%
		Upper		5	23	24	23.5	23.6	0.2%	589	349	-40.7%
	Middle	Lower		17	19	60	29.8	43.8	47.1%	2,241	2,215	-1.1%
		Upper		11	17	22	20.2	22.0	8.8%	668	1,531	129.3%
	Large	Regular		17	18	25	20.9	21.4	2.2%	704	575	-18.3%
	Luxury	Lower		11	17	27	20.3	23.3	14.6%	294	239	-18.6%
		Middle		6	15	19	17.4	17.9	3.3%	96	57	-41.2%
		Upper		10	14	23	19.1	19.9	4.4%	93	66	-29.3%
	Specialty	Sport		5	23	32	25.4	27.7	9.2%	179	165	-8.0%
		Small		9	18	25	20.3	21.3	4.8%	275	169	-38.7%
		Middle		9	18	25	20.3	21.3	4.8%	275	169	-38.7%
		Large		4	16	19	18.1	18.5	2.0%	32	15	-52.7%
		Luxury		7	21	26	22.6	24.7	9.1%	406	134	-67.0%
UV		CUV		17	17	36	24.7	35.0	41.6%	1,355	2,164	59.7%
		Middle		8	17	20	18.1	19.0	4.7%	346	19	-94.6%
		Large		10	17	32	22.8	32.0	40.3%	311	383	23.0%
		Mid. Luxury		5	16	21	17.8	20.9	17.4%	144	13	-90.9%
		Large Luxury		4	18	19	18.1	18.1	0.4%	311	278	-10.7%
	SUV	Small		14	15	19	16.7	18.4	10.1%	931	688	-26.1%
		Middle		9	13	16	15.3	15.8	3.7%	592	202	-65.9%
		Large		4	14	15	14.5	14.8	2.0%	60	13	-78.4%
		Mid. Luxury		7	13	14	13.1	13.4	1.9%	113	11	-90.2%
		Large Luxury		10	16	24	20.5	22.1	8.0%	543	881	62.1%
Truck/Van		Pickup		9	14	19	16.4	16.7	1.7%	2,183	1,424	-34.8%
		Small		12	17	20	18.8	19.2	1.9%	800	729	-8.9%
		Large		2	14	14	14.0	14.0	0.0%	139	32	-76.6%
		Small		2	14	14	14.0	14.0	0.0%	139	32	-76.6%
		Large		2	14	14	14.0	14.0	0.0%	139	32	-76.6%
		Luxury										