NBER WORKING PAPER SERIES

AUTOMOBILE PRICES IN MARKET EQUILIBRIUM: PART I AND II

Steven Berry

James Levinsohn

Ariel Pakes

Working Paper No. 4264

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1993

NBER Working Paper #4264
January 1993

AUTOMOBILE PRICES IN MARKET EQUILIBRIUM: PART I AND II

ABSTRACT

This paper develops new techniques for empirically analyzing demand and supply in differentiated products markets and then applies these techniques to analyze equilibrium in the U.S. automobile industry. Our primary goal is to present a framework which enables one to obtain estimates of demand and cost parameters for a broad class of oligopolistic differentiated products markets. These estimates can be obtained using only widely available product-level and aggregate consumer-level data, and they are consistent with a structural model of equilibrium in an oligopolistic industry. When we apply the techniques developed here to the U.S. automobile market, we obtain cost and demand parameters for (essentially) all models marketed over a twenty year period.

Steven Berry
Department of Economics
Yale University
37 Hillhouse Avenue
New Haven, CT 06520
and NBER

Ariel Pakes
Department of Economics
Yale University
37 Hillhouse Avenue
New Haven, CT 06520
and NBER

James Levinsohn
Department of Economics
University of Michigan
Ann Arbor, MI 48109
and NBER

Automobile Prices in Market Equilibrium: Part I.

Steven Berry
Yale University
National Bureau of Economic Research
James Levinsohn
University of Michigan
National Bureau of Economic Research
and
Ariel Pakes
Yale University
National Bureau of Economic Research

1. Introduction

This paper develops new techniques for empirically analyzing demand and supply in differentiated products markets and then applies these techniques to analyze equilibrium in the U.S. automobile industry. Our primary goal is to present a framework which enables one to obtain estimates of demand and cost parameters for a broad class of oligopolistic differentiated products markets. These estimates can be obtained using only widely available product-level and aggregate consumer-level data, and they are consistent with a structural model of equilibrium in an oligopolistic industry. When we apply the techniques developed here to the U.S. automobile market, we obtain cost and demand parameters for (essentially) all models marketed over a twenty year period. On the cost side, we estimate cost as a function of product characteristics. On the demand side, we estimate own- and cross-price elasticities as well as elasticities of demand with respect to vehicle attributes (such as weight or fuel efficiency.) These elasticities play crucial roles in the analysis of the likely effects of various policies. Further, these elasticities, together with estimated cost-side parameters, enable a more detailed economic analysis of the substantial changes that have occurred in the U.S. automobile industry over the period studied.

We would like to thank Don Andrews, Tim Bresnahan, Gary Chamberlin, Zvi Griliches, Jerry Hausman, Whitney Newey, and Frank Wolak for helpful comments. While working on this paper, Berry was an Olin Fellow at the NBER and Levinsohn was a National Fellow at the Hoover Institution, Stanford University. Each thanks their respective hosts. We gratefully acknowledge funding from National Science Foundation Grants SES-8821722 (to Richard Erikson and Ariel Pakes) and SES-9122672.

Our general approach posits a simple model which generates a distribution of consumer preferences over products. This model is then explicitly aggregated into a market-level demand system which, in turn, is combined with an assumption on cost functions and on pricing behavior to generate equilibrium prices and quantities. The primitives to be estimated are parameters describing the firms' marginal costs and the distribution of consumer tastes. The distribution of tastes determine elasticities, and these, together with marginal cost and a Nash assumption, determine equilibrium prices.

A familiar alternative to starting with a consumer-level utility function is to posit a simple functional form for the market level demand system. This requires some aggregation over products, since, for example, a constant elasticity demand system for, say, 100 products, would require estimating 10,000 elasticities. One might attempt to alleviate this difficulty by aggregating products *and/or* imposing constraints on elasticities, but, even apart from the appropriateness of the implied restrictions, aggregation methods which might seem useful for one policy experiment are unlikely to be useful for another. For example, an applied researcher investigating tariffs might be tempted to aggregate all foreign and all domestic cars. However the resulting model is unlikely to prove useful when investigating domestic competition or pollution taxes. There are further problems associated with the market-level demand approach. For example, this approach is poorly suited to analyze the likely effects of the introduction of new products. Also, as shown below, when one adopts a consumer-level approach, there are natural ways of incorporating the additional information provided by frequently available micro-level data.

One alternative to a demand system in the number of products marketed, an alternative that has been used extensively in the recent Industrial Organization literature, is a system which represents consumer preferences over products as a function of individual characteristics and of the attributes of those products. Advances in the discrete choice literature over the last two decades have generated much of the econometric methodology needed to use micro level data to estimate the parameters determining individual demands from this characteristics approach (see, for example, McFadden (1973) and the literature he cites in his 1986 review article). Moreover a few studies have, by using convenient (though quite restrictive) functional forms and distributional assumptions, been able to aggregate the individual demands generated by this approach into a market level demand system (see, for example, Morrison and Winston (1986)). Finally, attempts have been made to pair simple demand systems obtained in this manner with oligopolistic price setting models in a way that allows one to use the aggregate data to estimate the parameters needed to jointly determine equilibrium quantities sold and their prices (see Bresnahan (1987)).

We follow in this tradition, consider two problems that arise quite naturally in this framework, and provide computationally tractable methods for solving them. The first of the two problems concerns the imposed functional form of utility and the resulting pattern of cross-price elasticities (for a discussion, see Feenstra and Levinsohn (1991)). We show that to obtain substitution patterns that are consistent with generally accepted notions of substitution patterns for differentiated product markets, we need to be careful in modelling the interaction between consumer and product characteristics. The second problem addresses the correlation between prices, which are observed by the econometrician, and product characteristics which are observed by the consumer but not by the econometrician, and the bias in estimated elasticities that this induces. This is just the differentiated products analog of the traditional simultaneous equations problem in homogeneous product markets (the classic reference being Working (1926)). The resulting estimation strategy involves solving an aggregation problem in moving from the individual to aggregate demands (solved via simulation, as suggested by Pakes (1986)), and solving a non-linear simultaneous equations problem to account for endogenous prices (solved via an inversion routine as suggested by Berry (1991)).

Because we rely on mostly aggregate data, we do not have the very large number of degrees of freedom associated with more micro-level studies. This naturally raises concerns about obtaining precise estimates of the parameters of interest. We have two suggestions for ameliorating any precision problems that may arise. First we suggest using widely available data on the distribution of consumer characteristics to augment the market level information generally used in estimation, and provide some simple ways of incorporating this distributional information into our algorithm. Second, we use the recent literature on efficient instrumental variable estimators (see Chamberlain (1986)) to suggest instruments for our system. In models (such as ours) that involve Nash equilibria, efficient instruments will typically depend on a product's own characteristics as well as those of competing products. As a result the form of the optimal instrument function is of some independent economic interest. Since those instruments are difficult to compute, we follow Newey (1990) in using semiparametric approximations to them. Despite the fact that the efficient instruments depend upon the exogenous variables of all the competitors, a result on approximating exchangeable functions from Pakes (1992) allows us to find an easy to compute, manageable basis for this approximation (a fact which may be generally useful for estimating models that involve Nash equilibria).

The general framework used here is based upon: i) a joint distribution of consumer characteristics and product attributes which determines preferences over the products marketed; ii) price taking assumptions on the part of consumers; and iii) Nash equilibrium assumptions on the part of producers. This a very rich framework which we have not fully exploited. In particular, it is rich enough to incorporate nontrivial dynamics– dynamics which would endogenize both the distribution of product attributes and of consumer characteristics. Though we discuss this (and other) extensions in section 7 below, we have, for the purposes of this paper, stopped with a simple static model. This is a choice which limits both the richness and the usefulness of our results, and we intend to rectify it in future work. However it does allow us to focus in on the issues inherent in the two problems we do address: allowing for interactions between consumer and product characteristics, and allowing for unobserved (or unmeasured) product characteristics which induce a simultaneity problem in prices.

Allowing for richer preference patterns and unmeasured characteristics generates quite dramatic implications for own and cross-price elasticities (and for substitution patterns more generally). In particular, estimates of models which do not allow for interactions between consumer and product characteristics must *always* have the property that products with similar market shares have similar cross-price elasticities with respect to any third product. In the auto case, if a BMW and a Yugo have the same market share, they must of necessity have the same estimated cross-price elasticity of demand with respect to a Hyundai. Furthermore, models with similar market shares will always have similar estimated own-price elasticities. Alternatively, suppose we neglected unmeasured characteristics that in fact helped determine car choices. Then if, as our model suggests, these unobserved characteristics are positively correlated with price, this will tend to yield estimated own-price elasticities that are biased toward zero. Such biases are particularly problematic in an oligopoly context for they imply that mark-ups will be too high. In particular, inelastic estimated elasticities imply infinite markups or negative marginal costs, and this wreaks havoc on the pricing equation.

The Automobile Industry.

Few industries have been studied as intensively as the auto industry and with good reason. The market is a very important one. With sales topping $150 billion in 1989, the market is one of the largest in the U.S. and has ramifications for entire state economies. Moreover it is often at the heart of policy debates (in fields as diverse as international trade and environmental regulation) and it is a market which has evolved in important ways since 1973. Consequently, economists have tried to empirically analyze the U.S. car market for over 30 years.

Early work treated autos as a homogeneous product and estimated aggregate demand (*i.e.* Suits , 1958,). Griliches (1971) and later work by Ohta and Griliches (1976) adopted the hedonic approach. Their work was among the first to consider the automobile market at the level of the individual product, a feature which set the tone for much future research (examples include, Berkovec and Rust, 1985, Toder and Cardell, 1975, and Levinsohn, 1988). None of these studies gave much consideration to the production side of the model.

Perhaps the first attempt at simultaneously modeling and estimating the demand and oligopoly pricing sides of the market was Bresnahan's (1981) study. In that paper, Bresnahan adopts a Hotelling set-up and assumes a uniform density of consumers over the quality line. Feenstra and Levinsohn (1991) extend Bresnahan's work and allow products to be differentiated in multiple dimensions, but retain his assumption of the uniform density of consumers. Manski (1983) investigates the (perfectly competitive) supply side and demand side of the Israeli automobile market. By allowing for products that are differentiated in multiple dimensions, richer distributions of taste parameters, and unobserved (to the econometrician) product characteristics, and then showing how to provide consistent estimates of models with all these features, we integrate and extend the advances in this literature, thereby taking a step towards a more detailed understanding of behavior in the auto market.

A Road Map.

The remainder of the paper is organized as follows. The next two sections describe our theoretical model. Section 2 contains a discussion of utility and demand, while section 3 models firm behavior and derives industry equilibrium. Section 4 presents the estimators and describes their properties while Section 5 provides the required computational techniques. The data and estimation results are discussed in section 6. This section also provides a quick review of alternative models and compares the estimates from our models to those of some alternatives. We conclude and discuss possible extensions in Section 7. Three appendices discuss details used in Sections 4 and 5.

2. Theory: Utility and Demand.

Our demand system is derived from a standard discrete choice model of consumer behavior; a model that makes the choices of consumers a function of their characteristics and the characteristics of the available products. The market demand for a given product is obtained by explicitly aggregating over individual choices. (For background on demand systems obtained in this manner see McFadden (1981) and the literature cited there as well as the product differentiation literature cited in Shaked and Sutton 1982), Sattinger (1984), Perloff and Salop (1985), Bresnahan (1987), and Anderson DePalma and Thisse, (1989) (among others)) We then combine this demand system with a cost function, and embed these two primitives into a model of price setting behavior in differentiated products markets. The demand and pricing equations that this model generates give us the system of equations that we take to the data.

Most of this paper assumes that we do not have data that matches individual characteristics to the products those individuals purchased. Consequently we proceed (as does much of the prior literature on the empirical analysis of equilibrium in markets for differentiated products[1]) by considering the problem of estimating all the parameters of the

[1] For examples see, Bresnahan (1987), and Feenstra and Levinsohn (1991.)

demand system from product level data (*i.e.* from information on prices, quantities, and the measurable characteristics of the products). We then extend the discussion to allow for the possibility of incorporating exogenous (and frequently available) information on the *distribution* of individual characteristics (e.g. the distribution of income and/or family size). Only later do we come back to the advantages of having data that matches consumer characteristics to the products those consumers purchased.

Our specification posits that the level of utility that a consumer derives from a given product is a function of both a vector of individual characteristics, say ν, and a vector of product characteristics, say (x, ζ, p). Here p represents the price of the product, and x and ζ are, respectively, observed and unobserved (by the econometrician) product attributes. That is, the utility derived by consumer i from consuming product j is given by the scalar value

$$U(\nu_i, p_j, x_j, \zeta_j; \theta),$$

where θ is a k-vector of parameters to be estimated.[2] Consumers with different ν make different choices, and to derive the aggregate demand system we integrate out the choice function over the distribution of ν in the population of interest.

In this context we note that throughout we will take ν to have a *known* distribution. This distribution may either be the empirical distribution of the characteristic of interest, or a standardized distribution whose standardization parameters are estimated (unit normals for example, in which case the standardization parameters are the associated mean vector and covariance matrix). Thus it should be understood that θ includes the standardization

[2] We should note that it is not necessary, indeed not even traditional, to take $U(i,j)$, the utility consumer i derives from consuming product j, as the primitive of the problem. We could, for example, have begun with a choice problem over a vector of continuous, as well as over our discrete, products, and then derive $U(i,j)$ as the utility from choosing alternative "j" given that the consumer does the best he or she can over the continuous products. Were we to have started out at the more primitive level, and then derived our $U(i,j)$, we would have added notation, without changing the substantive discussion of this paper. We should note, however, that when there is more than one discrete choice to be made (including, possibly, the choice of whether to buy a second unit of the product of interest), the reduction to the type of model we are using would, in general, require us to let the index j run over all possible combinations of discrete choices. That is, without further constraints the choice set would grow exponentially in the dimension of the discrete choice vector, and this will increase the computational burden of the problem significantly. In this case, it may be worthwhile to go back to the more primitive problem and try to use it to eliminate possibilities from the choice set.

parameters we wish to estimate, as well as parameters that describe the utility surface conditional on an individual's characteristics.

Consumer i chooses good j if and only if

$$U(\nu_i, p_j, x_j, \zeta_j; \theta) \geq U(\nu_i, p_r, x_r, \zeta_r; \theta), \text{ for } \; r = 0, 1, ..., J,$$

where alternatives $r = 1, ..., J$ represent purchases of the competing differentiated products. Alternative zero, or the outside alternative, represents the option of not purchasing any of those products (and allocating all expenditures to other commodities). It is the presence of this outside alternative that allows us to model changes in the total quantity of automobile purchases as a function of the prices and characteristics of the cars marketed (*i.e.* that allows us to obtain a nondegenerate aggregate demand function).

Let

$$(2.1)\ A_j = \{\nu : U(\nu, p_j, x_j, \zeta_j; \theta) \geq U(\nu, p_r, x_r, \zeta_r; \theta), \text{ for } r = 0, 1, ..., N\}.$$

That is A_j is the set of values for ν that induces the choice of good "j". Then, assuming ties occur with zero probability, and that $P_0(d\nu)$ provides the density of ν in the population of interest, the market share of good "j" as a function of the characteristics of all the goods competing in the market is given by

$$(2.2a)\ s_j(p, x, \zeta; \theta) = \int_{\nu \epsilon A_j} P_0(d\nu)\ .$$

Let the J-element vector of functions whose "j^{th}" component is given by (2.2a) be $s(\cdot)$ where

$$(2.2b)\ s(p, x, \zeta; \theta) = [s_1(p, x, \zeta; \theta), s_2(p, x, \zeta; \theta), ..., s_J(p, x, \zeta; \theta)]'\ .$$

Thus if M is the number of consumers in the market of interest, the vector of demands for the J products is

$$Ms(p, x, \zeta; \theta).$$

(In the empirical work below we set M equal to the number of households in the U.S. economy, though in general it could play the role of a parameter to be estimated.)

Subsection 2.1 begins with a discussion of the implications of alternative functional forms for the utility function, and then considers integrating exogenous information on the distribution of consumer characteristics into the analysis.

2.1 *Functional Forms and Substitution Patterns.*

A special case of the model in (2.1) and (2.2) is

$$(2.3)\ U(\nu_i, p_j, x_j, \zeta_j; \theta) = x_j\beta - \alpha p_j + f(\nu_i, \zeta_j) \equiv x_j\beta - \alpha p_j + \xi_j + \epsilon_{i,j} \equiv \delta_j + \epsilon_{i,j},$$

where

$$\delta_j = x_j\beta - \alpha p_j + \xi_j.$$

Without unobserved characteristics, i.e. $\xi_j = 0$, this specification has been used extensively in the empirical literature. The mean of the ϵ vector in the population of consuming units is assumed to be zero so that for each j, ξ_j is the mean of the unobserved component [of $f(\nu_i, \zeta_j)$], and δ_j is the mean of $U_{i,j}$. Importantly, it is also assumed that differences in the distribution of the $\epsilon_{i,j}$ across j are independent of the observed characteristics of the products (of the x_j). Indeed, most often it is assumed that $[\epsilon_{i,0}, ..., \epsilon_{i,J}]$ is a vector of independently and identically distributed random variables.

It is easy to see why (2.3) together with an *i.i.d.* assumption on the distribution of the ϵ has become a popular specification. This specification makes it particularly easy to compute the model's implications for the observed market shares conditional on alternative possible values for the parameter vector (that is, to compute the integral in 2.2). In particular, these assumptions imply that

$$(2.4)\ s_j = \int_\epsilon \Pi_{q \neq j} P(\delta_j - \delta_q + \epsilon) P(\epsilon)\ .$$

(2.4) shows that computation of the implied market share requires, at most, computation of a unidimensional integral. Of course, if the ϵ are distributed multivariate extreme value there is a closed form for 2.4 (see below) and there is no need to compute any integral in order to obtain the model's implications for the market shares, or the s_j's.

Despite this computational simplicity the assumption that the utility function is *additively separable* into an effect of the product's characteristics that is the same for all consumers (the δ_j in 2.3), and an effect of the consumer characteristics that is independent of the observed product characteristics (the $\epsilon_{i,j}$ in 2.3) is problematic.[3] In particular additive separability almost always has counter-intuitive behavioral implications. Thus, in our auto example, we expect the utility generated from different sized cars (a product characteristic) to depend on family size (a characteristic of the consuming unit), and the effect of price to depend on income.

For present purposes, the important point is that the same logic that leads us to question the intuitive basis for the additively separable specification in (3), generates a set of substitution patterns, and hence a set of (cross and own) price derivatives, as well as a set of responses to the introduction of new products, that *can not* possess many of the features that we are quite sure lie behind actual substitution patterns. Moreover, any specification that misrepresents substitution patterns will misrepresent the nature of the competitive interactions among firms, including the implications of alternative behavioral or institutional assumptions on the likelihood of alternative outcomes.

One way to see the problem with the specification in (2.3) is to note that it generates substitution effects which depend only on the vector of δ_j indices. Since, under mild regularity conditions (see Berry (1992)), there is a unique vector of market shares associated with each vector of δ-indices, an implication of the specification in (2.3) is that the cross-price elasticities between any two products, or, for that matter, the similarity in their price and demand responses to the introduction of a new third product, depends only on their market shares. That is, conditional on market shares, substitution patterns do not depend on the observable characteristics of the product.

Thus if we were using the specification in (2.3) to analyze an automobile market in which a relatively inexpensive Yugo and a relatively expensive Mercedes had the same market shares, then the parameter estimates would have to imply that the two cars have the same cross-price derivative with respect to *any* third car. In particular, the model

[3] (2.4) assumes more than additive separability. It also assumes that $\delta(j)$ is a linear function of product characteristics, and that the distributions of the $\epsilon_{(}i,j)$ is identical across "j"; but these assumptions are primarily for expositional simplicity. They can be relaxed with only minor modifications to the discussion that follows (see below). For early discussions of the implications of related specifications on aggregate demand patterns see the appendix to Hausman (1978), McFadden (1981), and Schmalensee (1985).

would necessarily predict that a reduction in the price of a BMW would generate equal reductions in the demand for Yugos and for Mercedes. This contradicts the intuition which suggests that couples of goods whose characteristics are more "similar" should have higher cross-price elasticities. That is when the price of a large car, say car A, goes up we expect the demand for other large cars to go up disproportionately.

The reason we expect this to happen is that the consumers who would have chosen car A at the old prices, but now may not, are consumers who have a preference for large cars. If they move to a different car, they are likely to move to a different car which is also large. Similarly, when a new car enters the market, we expect it to have a large effect on the demand for cars with similar characteristics.

For analogous reasons, the specification in (2.3) implies that two products with the same market share will have the same own-price demand derivatives. For example, if a Jaguar and a Yugo have the same market share, the specification in (2.3) implies that they must have the same own-price derivative. In an oligopoly context, this is troubling for it implies that the two products must have the same markup over marginal cost (at least assuming that firms all sell only a single product, for the extension to multi product firms see below). Intuitively, however, we expect markups to be determined by more than market shares. They ought also to be determined by the number of competing products which are "close" in product space, and, because consumers who buy more expensive goods are likely to have lower marginal utilities of income, by the price of the product.

Note that the counterintutive implications of (2.3) are solely a result of the additive separability embodied in that specification. In particular, they do not depend on the particular distributional assumption made on the $\epsilon_{i,j}$. The *i.i.d.* assumption on the $\epsilon's$ alone is enough to imply that a consumer that is induced to substitute away from any given choice will make choices that resemble the choices of "an average" consumer; regardless of his or her original choice. That is, the consumer who substitutes away from product j as a result of the increase in the price of that product will tend to substitute toward other popular products, not to other similar products.[4]

[4] If $U(i,j) = g[p(j), x(j), \zeta(j); \theta] + \epsilon_{i,j}$, that is, if utility was not linear in the attributes of the product but was additively separable into a component which depended on observable product characteristics and

We now consider alternative ways of allowing for interaction between individual and product characteristics. A familiar starting point is to allow each individual to have a different preference for each different observable characteristic. This generates the traditional random coefficients model

$$(2.5)\ U(\nu_i, p_j, x_j, \zeta_j; \theta) = x_j\bar{\beta} - \alpha p_j + \xi_j + \textstyle\sum_k \sigma_k x_{jk}\nu_{ik} + \epsilon_{ij},$$

where $(\nu_i, \epsilon_i) = (\nu_{i,1}, \nu_{i,2}, ..., \nu_{i,K}, \epsilon_{i,0}, \epsilon_{i,1}, ..., \epsilon_{i,N})$ is a mean zero vector of random variables with (a known) distribution function (P_0). Now the contribution of x_k units of the k^{th} product characteristic to the utility of individual i is $(\bar{\beta}_k + \sigma_k\nu_{i,k})x_k$, which varies over consumers with $\nu_{i,k}$. We scale ν_{ik} such that $E(\nu_{ik}^2) = 1$, so that the mean and the variance of the marginal utilities associated with characteristic k are $\bar{\beta}_k$, and σ_k^2 respectively.[5] As noted below, this specification is particularly tractable if ϵ_i consists of $i.i.d$ extreme value deviates, and ν_i has a multivariate normal distribution.

The distribution of the utility obtained from consuming good j can still be decomposed into a mean

$$\delta_j = x_j\bar{\beta} - \alpha p_j + \xi_j$$

and a deviation from that mean

$$\mu_{ij} = \sum_k \sigma_k x_{jk}\nu_{ik} + \varepsilon_{ij},$$

but now the properties of the distribution of the deviation from the mean depends on the interaction between consumer preferences for different characteristics and the characteristics of the product. As a result consumers who have a preference for size will tend to

a component which depended on individual characteristics, then the elasticity of demand of product A with respect to the price of product B might depend on the characteristics of product A, but it would not depend on the relationship of the x-vectors of the two products. Alternatively, if the distribution of the $\epsilon(i,j)$ were different for different products but independent of "$x(j)$" then the substitution effects could depend on both the mean level of utilities and on the appropriate distribution function, but it would not depend on the relationship of the x-vectors of the two products.

[5] We will assume that the distribution of the $[\nu(i,1),,..,\nu(i,K)]$ factors into a product of independent densities. This is for expositional convenience, with the addition of some notation we could easily allow for patterns of correlation among them.

attach high utility to all large cars, and this will induce large substitution effects between goods that have similar characteristics.

Now if (ϵ_i, ν_i) consists of mutually uncorrelated random variables, then

$$(\mu_{i,0}, \mu_{i,1}, ..., \mu_{i,J}),$$

distributes across individuals as a vector of correlated, heteroskedastic, random variables with

$$Var(\mu_{i,j}) = \Sigma_k[\sigma_k^2 x_{j,k}^2 + \sigma_{\epsilon(k)}^2], \text{ and } \quad Cov(\mu_{i,j}, \mu_{i,q}) = \Sigma_k \sigma_k^2 x_{j,k} x_{q,k}.$$

One can reinterpret the x_{jk} in the above formula as deviations from the mean value of that characteristic over all products. Hence, the formula above states that the utility attached to a given good will vary a great deal across consumers if the good has deviant values for characteristics with large $\sigma's$ (characteristics whose values differ greatly among consumers). Analogously, the utility associated with goods with similar values for characteristics whose values differ greatly among consumers will be highly positively correlated.

The second point to note here is that now the integral required to calculate the model's implications for market shares conditional on alternative possible values of the parameter vector (the integral in 2.2 above) can no longer be reduced to a simple unidimensional integral (as in 2.4), but rather is a $K+1$ dimensional integral. We show below that if the ϵ vector consists of independent extreme value deviates, the integral required to calculate the market shares can be expressed as a K dimensional integral and further computational simplifications are possible; but that still leaves us with a high dimensional integral to compute. We solve this computational problem below via aggregation by simulation, a technique introduced by Pakes (1986).

Before leaving the specification in (2.5) we consider its relationship to more general approximations to our utility function, $U(\nu_i, p_j, x_j, \zeta_j; \theta)$. One can get to the random coefficients specification in (2.5) from our underlying utility function by assuming that there are K unobserved individual characteristics that determine the marginal utilities of each product, and an unspecified number of individual specific characteristics that interact with an unspecified number of product characteristics, as in (2.3), to produce the $\xi_j + \epsilon_{i,j}$. This can be viewed as a restricted second order approximation to the utility function. It is restricted in that it does not allow for any interaction between individual characteristics and price, does not allow for a full set of second order coefficients (as that would require squared terms in the ν and the x), and assumes interactions between the ξ and the individual characteristics that produce an ϵ whose distribution is independent of ξ. We can relax each of these restrictions at the cost of an additional computational burden. (The additional burden would be the highest were we to relax the third restriction.) Moreover this line of reasoning could be pushed further by adding higher order terms in an attempt to get a more flexible approximation to the underlying utility surface.

The basic problem we see with this route is that the dimensionality of the ν vector, and the way we constrain its interactions with the observable characteristics of the product, are set arbitrarily without bringing to bear any extra information we might have on the problem of interest. We often have a fair amount of prior information on both which characteristics of consumers determine preferences over different product characteristics, and on the functional form of the interaction between the consumer and the product characteristics. Moreover the distribution of at least some of the consumer characteristics of interest is often either known, or can be estimated from another data source, and that information can also be brought to bear on the estimation problem.

We now modify our strategy to make more intensive use of these alternative sources of information. In particular we assume we can specify (at least some of) the consumer characteristics which cause differences in household's evaluations, and then make use of the available information on the distribution of those characteristics in the population of interest. We could do this in the context of the pure random coefficients model discussed above (e.g. assume that the distribution of the random taste parameter on car size depends on family size, and make use of the observed distribution of family size in the market of interest). However, once we allow income to be a relevant consumer characteristic then incorporating it into this random coefficients specification in a way that allows for both sensible interactions with price, and declining marginal utility of income, might require us to use a polynomial approximation of a high order. As a result we nest the random coefficients framework into a more traditional utility function, one which incorporates diminishing marginal utility and the interaction of income and price in a parsimonious way, and then use the resultant functional form to structure estimation.

Perhaps the most familiar starting place here is a Cobb-Douglas utility function in expenditures on other goods and services, and characteristics of the good purchased

$$(2.6)\ U(\nu_i, p_j, x_j, \zeta_j; \theta) = (y_i - p_j)^{\alpha} G(x_j, \xi_j, \nu_i) e^{\epsilon(i,j)},$$

where y is income, ϵ provides the effect of the interactions of unobserved product and individual characteristics, and the form of $G(\cdot)$ as well as the distribution of ν, are to be determined by the problem at hand.

In our empirical example we assume that $G(\cdot)$ is linear in logs and has the random coefficient specification discussed above, so that if $u_{ij} = \log[U_{ij}]$, then

$$(2.7a)\ u_{ij} = \alpha log(y_i - p_j) + x_j\bar{\beta} + \xi_j + \Sigma_k \sigma_k x_{jk} \nu_{ik}\ ,$$

for $j = 1, ..., J$, while

$$(2.7b)\ u_{i0} = \alpha log(y_i) + \xi_0 + \sigma_0 \nu_{i0} + \epsilon_{i0}.$$

Two points should be noted about this specification. First, our current data set does not have information on differences in the value of the outside alternative across consumers (differences that would be generated by, among other diverse factors, differences in access to public transportation and differences in used car holdings). Thus, to account for the possibility that there is more unobserved variance in the idiosyncratic component for the outside then for the inside alternatives, we have allowed for an extra unobserved term in the determination of u_{i0} (the ν_{i0}).[6] Second, the consumer characteristic terms that interact with product characteristics are now denoted by

$$\nu_i = (y_i, \nu_{i0}, \nu_{i1}, ..., \nu_{iK}).$$

[6] Note that since market shares depend only on differences in utilities, the actual estimation algorithm ends up subtracting the $u(i,0)$ in (2.7b) from the $u(i,j)$, and estimating a model where the outside alternative is "normalized" to zero. Given the specification in (2.7b), this implies that there is a random coefficient on the constant terms in the mean utility specification for the inside goods. We should also note that much of our current modelling and data gathering activity is directed at enabling us to incorporate more detailed information on the outside alternative (see the discussion in the extensions section below).

We have used special notation for income here both because it enters the utility function in a special way, and because it is a variable whose (marginal) distribution can be obtained from the March Current Population Survey (at least up to an error process). As a result, provided one is willing to assume a parametric form for the distribution of $(\nu_{i1}, ..., \nu_{iK})$ conditional on y_i, we can use the CPS to determine the distribution of y in our population, and save on the parameters that need to be estimated inside our algorithm.

There are two characteristics of (2.7) that are central to the rest of this paper. First, it allows for interactions between consumer and product characteristics, and second, it allows us to make use of exogenous data on the distribution of income (and quite possibly the joint distribution of income and other household characteristics) in a natural and parsimonious way. The first characteristic is important because it enables the model to generate reasonable substitution patterns, the second allows us to get more precise parameter estimates (and precision becomes more of a concern when we allow our specification to be flexible enough to generate reasonable substitution patterns).

2.2 *Endogenous Prices*

This paper assumes that equilibrium in our differentiated products market is Nash in prices. Hence, producers choose prices to maximize their profits conditional on the prices and characteristics of their competitors. Consequently, if producers know the values of the unobserved characteristics (of the $\xi's$) even though we as econometricians do not, then prices will be a function of them (as well as of observable characteristics). This generates a differentiated products analog to the classic simultaneity problem in the analysis of demand and supply in homogeneous product markets. (The usual reference here is Working (1926), while for a history of the econometrics of demand and supply analysis in homogeneous product markets see Morgan (1990, chapter 2)). The simultaneity problem is complicated by the discreteness in the choice set of individuals which generates individual demand functions that are a nonlinear function of the attributes (in particular of the unobserved attributes) of the product. This in turn makes aggregate demand a nonlinear function of these product characteristics. Berry (1992) suggests one approach to obtaining consistent estimates of parameters of the demand system, and proves its viability under certain functional form restrictions. This subsection begins by discussing the importance of unobserved demand characteristics and the resulting endogeneity of prices, and then reviews and extends Berry's approach for estimating models which contain them. Later sections consider more detailed aspects of the estimation problem (including the question of optimal instruments), and the computational techniques required to implement them (including ways of simplifying the computation of the optimal instrument vector).

All prior empirical work on discrete choice models of demand which estimated price effects has specified that the unobserved component in the function determining the utility of each alternative is mean zero and independent across agents.[7] This is true regardless of whether that prior work was on micro data (data which matched individual characteristics to individual choices), or on more aggregate market data (which matched market shares to the price and attributes of the product). The contrast between this specification of the properties of the disturbances, and the specification used in the more traditional homogeneous product models of demand, is most transparent in the aggregate implications of the discrete choice studies. If the disturbance vector is mean zero and independently distributed across agents, then the aggregation process integrates it out, leaving aggregate market shares a function of only observables and the parameters to be estimated. Moreover since aggregate shares do not depend on any disturbance, there can be no simultaneity problem in the demand equation.

In contrast, aggregate demand in homogeneous product markets is typically specified to have a nonzero disturbance associated with it– a disturbance which is generally explained in terms of unobserved determinants of demand that are correlated across agents in a given market (be that market a region or a time period). It is the presence of this disturbance in the aggregate demand equation in homogeneous product markets that induces the simultaneity problem that has justly received so much attention in the empirical analysis of demand and supply. If these disturbances are known to consumers participating in the market (and if demand depends upon them, one would expect this condition to be satisfied), and if there is any equilibrating mechanism in the market at all, then equilibrium

[7] One exception is Berry's (1991) study of airline hubbing, which includes an aggregate market-specific demand error which is correlated with prices. However, that paper uses a very restrictive functional form for utility and an estimation procedure which is not robust to non-uniqueness of equilibria.

quantities and prices will depend upon the disturbances. This, in turn, generates a correlation between the disturbance and the price variable and the need for alternatives to ordinary least squares estimation techniques.[8]

By assuming away any dependence in the distribution of disturbances associated with a given product across agents, the discrete choice literature not only ignores potential simultaneity problems, but it also generates a rather embarrassing "over fitting" phenomena. If there is no "structural" disturbance in the market share equation, then the only source of error which can result in differences between the predictions of the model and the data is sampling error. Sampling error results from the fact that we do not have data on the share of the entire population that purchased a particular good. (Alternatively, the individuals in the market are assumed to be a random draw from some larger superpopulation, and the observed market shares fluctuate about the shares that would be purchased by the superpopulation.) As a result the data on the proportions of different goods purchased can differ from the underlying population proportions that the model predicts would be purchased. The sampling process, however, implies that the sample's proportions are distributed multinomially about the true population proportions. The variance-covariance of this distribution is proportional to one over the size of the sample. (It equals $N^{-1}\{\text{diag}(s) - ss'\}$, where s is the vector of market shares, N is the size of the sample, and $\text{diag}(x)$ is a diagonal matrix with x on the principal diagonal). For sample sizes as large as those typically found in aggregate studies, this variance is just too small to account for any noticeable discrepancy between the data and the model (so that the familiar χ^2 test for the adequacy of the model's restrictions on the multinomial proportions is rejected with probability close to one).[9]

All the utility specifications we introduced in the last subsection had disturbances with a product specific mean (the vector of $\xi's$). These product specific means are the analog of

[8] A rare exception occurs when we have both marginal cost pricing and an unobservable in the demand function which is mean independent of marginal cost. Marginal cost pricing is at most a limiting case in differentiated product markets, and the assumption that the unobserved quality of a product is not correlated with its marginal cost seems inappropriate, at least as an a priori restriction.

[9] Similar overfitting phenomena have been a source of concern in the biometrics literature for some time; see, for example, Haseman and Kupper, 1979, or Williams, 1982. Though they do not worry about simultaneity, their conceptual solution to the overfitting problem is similar to the one we shall use (allowing for unobserved determinants of the cell probabilities).

the disturbance in the demand system in homogeneous product markets, and are meant to capture the effect of unmeasured characteristics that lead to correlation in the intensity of preferences for a given product across consumers. In the automobile example, ξ reflects the impact of difficult to quantify aspects of style, prestige, reputation, and past experience that affect the demand for different products, as well as (indeed perhaps most importantly) the effects of quantifiable characteristics of the car that we simply do not have in our data. As one might expect, the introduction of these unobserved product characteristics will alleviate the overfitting problem noted above. However this is not the major reason for going to the trouble of allowing for the unobserved product characteristics. Our primary concern is that if any such unobserved product characteristics are important, and our data indicate that they are, prices will be correlated with them, and the estimates of price effects will be biased in known ways.[10] This is the precisely the same logic that leads to the belief that O.L.S. estimates of price effects in traditional homogeneous product demand systems are positively biased.

As in traditional homogeneous goods models that use aggregate data, we will assume that ξ is mean independent of the observable characteristics and derive estimators of the parameters of interest from the orthogonality conditions those assumptions imply.[11] The difference between our case and the homogeneous product case is that, because of the nature of the choice set in discrete choice models, the demand of a given individual, and hence the sum of the demand across individuals, becomes a nonlinear function of the ξ; *i.e.* $Q^d = Ms(x, \xi, p; \theta)$, where Q^d is the vector of quantities demanded. Consequently

10 Since in most discrete choice frameworks any unobserved characteristic will enter the moment condition defining the estimator in a nonlinear fashion, omitting it will always cause the parameter estimates to be inconsistent, regardless of whether the unobserved characteristic is a determinant of price. Formally then, the additional fact that our behavioral assumptions are likely to generate a positive correlation between the unobserved characteristic and price just helps us to sign the bias.

11 This is the identifying assumption that we will maintain throughout this paper, but it is not the only identifying assumption one could integrate with our framework, and the appropriate assumption is likely to depend on the market being analyzed and the nature of the data that is available. For example Das, Olley, and Pakes, (in process) use variation in the market shares of different products across income classes to obtain estimates of both the parameters describing the interaction between consumer and product characteristics, and pointwise estimates of the mean utility levels (our δ's) without imposing any restriction on the joint distribution of the ξ and the observable characteristics of the firm. Similar possibilities arise when we have regional information on market shares, or when we have micro data (see the extensions section of the paper).

the orthogonality between ξ and the vector of x characteristics cannot be used as a basis for estimation without first transforming the observed quantity, price, and characteristic data into a linear function of the ξ. Berry (1992) shows that, for any given value of the parameter vector and list of product attributes, there is, under mild restrictions on the relevant functional forms, a unique map from market shares back to the ξ. Below we show how to compute this "inverse map" and then combine it with orthogonality conditions between the computed ξ vector and the product attribute vector to obtain a method of moments estimator for θ.

This procedure only requires the assumptions required for consistency of instrumental variable estimators of demand parameters in homogeneous product markets. In particular we do not require an explicit assumption on the distribution of the ξ, just that they be mean independent of the observable attributes of the product. Furthermore, the procedure does not depend on the exact form of the pricing rule. On the other hand, since the pricing rule in a Nash price equilibrium depends on the true values of the demand parameters, joint estimation of the pricing and demand equations ought to increase the efficiency of the estimates of the demand parameters. We show how to derive the pricing equation and obtain estimates of the model's parameters from a system composed of the demand and pricing equations below. Note, however, that neither the estimation procedure for the demand equation, nor the procedure for the system composed of the demand and pricing equations, will require that the underlying game generate a unique equilibria (and it is difficult to prove that the solution to Nash pricing equilibria in differentiated product markets will be unique, see Caplin and Nalebuff (1991)). It is, however, the case that one cannot use our parameter estimates to compute the changes in prices or quantities that would result from a policy (or an environmental) change without a way of selecting among alternative equilibria (should they exist).

Note also that, again just as in the homogeneous products case, we will need to obtain an effective instrument for price. Of course, any factor which varies across products and effects costs but not demand will do. In addition, the fact that it is natural to model a differentiated products market as an oligopolistic or monopolistically competitive market implies that the characteristics of other products will be appropriate instruments for a product's price. This raises the question of how to obtain efficient instruments when any function of competitors' characteristics (as well as functions of the product's own characteristics) are potential instruments; a question which is sure to arise repeatedly in econometric models of non-perfectly competitive markets. We use results from the recent econometric literature on efficient estimators subject only to conditional moment restrictions (and mild regularity conditions; see Chamberlain (1986)) to explore this issue, a solution, and use a result on approximations to exchangeable functions in Pakes (forthcoming) to provide an easy way to compute it.

We cannot go on to a more detailed discussion of the estimation algorithm without first introducing the pricing equation.

3. Cost Functions and The Pricing Problem of the Multiproduct Firm.

We take as given that there are F firms, each of which produce some subset, say $\mathcal{J}_f$, of the J products. For simplicity we begin by assuming that the marginal cost of producing the goods marketed are both independent of output levels and log linear in a vector of cost characteristics. These assumptions are made only for expositional convenience and we relax them in our investigation of the robustness of our empirical results below.

The cost characteristics are decomposed into a subset which are observed by the econometrician, the vector w_j for model j, and an unobserved component, ω_j. Note that we might expect the observed product characteristics, the x_j, to be part of the w_j, and ω_j to be correlated with ξ_j. This because larger cars, or cars with a larger unobserved quality index, might be more costly to produce. This overlap is accounted for in our estimation algorithm.

Given these assumptions the marginal cost of good j , say mc_j, is written as

$$(3.1)\ ln(mc_j) = \mathrm{w}_j\gamma + \omega_j,$$

where γ is a vector of parameters to be estimated. The fact that we have assumed log marginal cost is additively separable into its observed and unobserved components allows us to suffice with assumptions on the conditional mean of ω in the estimation algorithm proposed below (in contrast to needing more detailed assumptions on its conditional distribution).

Given the demand system in (2.1) and (2.2) (derived from the utility function in, say, (2.7)), the profits of firm "f", say Π_f, are given by

$$(3.2)\ \Pi_f = \Sigma_{j\in\mathcal{J}_f}(p_j - mc_j)M\ s_j(p,x,\xi;\theta)\ ,$$

with mc_j given by (3.1) above. Each firm is assumed to choose prices for its products to maximize its profit given the attributes of its own products, and the prices and attributes of all competing products.

We assume that a Nash equilibrium to this pricing game exists, and that the equilibrium prices are in the interior of the firms' strategy sets (the positive orthant). While Caplin and Nalebuff (1991) provide a set of conditions for the existence of equilibrium in the case of single product firms, their theorems do not easily generalize to the multiproduct case. As a result we do not have an analytic proof of the existence of equilibrium in our model (*i.e.* a proof that would be valid for any value of the model's parameter vector). However, we will be able to check numerically whether our final estimates are consistent with the existence of an equilibrium. As noted, none of the properties of the estimates we derive require that there be a unique equilibrium associated with any given value of the parameter vector.

Given our assumptions, any product produced by firm f, or any $j\in\mathcal{J}_f$, must have a price, p_j, that satisfies the first order conditions

$$(3.3)\ s_j(p,x,\xi;\theta) - \sum_{r\in\mathcal{J}_f}(p_r - mc_r)\frac{\partial s_r(p,x,\xi;\theta)}{\partial p_j} = 0\ .$$

The J first-order conditions in (3.3) imply pricing equations, or price-cost markups $(p_j - mc_j)$ for each good. To obtain these, define a new J by J matrix, Δ, whose (j,r) element is given by:

$$(3.4)\quad \Delta_{jr} = \begin{cases} \frac{-\partial s_r}{\partial p_j}, & \text{if r and j are produced by the same firm;} \\ 0, & \text{otherwise.} \end{cases}$$

In vector notation the first order conditions can then be written as

$$s(p,x,\xi:\theta) - \Delta(p,x,\xi;\theta)[p - mc] = 0.$$

Solving for the price-cost markup gives

$$p = mc + \Delta(p, x, \xi; \theta)^{-1} s(p, x, \xi; \theta) \ .$$

Note that prices are additively separable in marginal cost and the markup over marginal cost,

$$(3.5)\ b(p, x, \xi;\ \theta) \equiv \Delta(p, x, \xi; \theta)^{-1} s(p, x, \xi; \theta).$$

The vector of markups in (3.5) depends only on the parameters of the demand system and the equilibrium price vector. However, since p is function of ω, $b(p, x, \xi;\ \theta)$ is a function of ω, and cannot be assumed to be uncorrelated with it (the fact that ξ is correlated with ω also generates a dependence between the markups and ω).

Substituting in the expression for marginal cost, we obtain the pricing equation we take to the data

$$(3.6)\ ln(p - b(p, x, \xi; \theta)) = \mathrm{w}\gamma + \omega \ .$$

Just as in estimating the demand equation, estimates of the parameters of (3.6) can be obtained from orthogonality conditions between ω and alternative functions of the vector of the products own, and its competitors, characteristics. Consequently, the discussion at the end of section 2.3 on efficient instruments is also directly relevant here. In addition, as noted above, the fact that the markup is determined by the demand parameters implies that it would increase the efficiency of the estimation algorithm if we estimated the parameters of the demand system jointly with the parameters of the pricing function in (3.5).

We turn to a more detailed discussion of this algorithm now, beginning with a description of the estimators (section 4), and then moving on to a discussion of ways to compute them (section 5). The reader who is not interested in these details should now be able to move directly to the discussion of the empirical results in section 6.

4. The Estimation Algorithm.

To keep the exposition simple, we begin by maintaining some simplifying assumptions that we later remove. In particular, although we will actually use panel data, we start by assuming that our data consists of a single cross section of the autos marketed in a given year. If J is the number of autos marketed, the data set then contains J vectors $(x_j, \mathrm{w}_j, p_j, q_j)$, and a number of households sampled, n, which, when combined with the information on purchases, can be used to compute the share of the outside alternative (the number of households that do not purchase any auto). Thus, the observed vector of sampled market shares, denoted s_n, belongs to the $J+1$ dimensional unit simplex. (There is a share for each good marketed and one for the outside alternative).

The assumptions on the data generating process are as follows. Market shares are calculated from the purchases of a random sample of n consumers from a population with a distribution of characteristics, ν, given by $P_0(\cdot)$. This population abides by the model's decision rules at $\theta = \theta_0$. Letting s_0 denotes the vector of shares in the underlying population, the multinomial sampling process implies that s_n converges to s_0 at rate $\sqrt{n}$, or $(s_n - s_0) = O_p(1/\sqrt{n})$. The $(\xi_j, \omega_j, x_j, \mathrm{w}_j)$ vectors that characterize the primitive product characteristics are exchangeable draws from some superpopulation of possible characteristic vectors. The distribution of these vectors in this superpopulation has the property that if $z_j = [x_j, \mathrm{w}_j]$, and $z = [z_1, ..., z_J]$ then

$$(4.1)\ E[\xi_j|z] = E[\omega_j|z] = 0, \text{ and } E[(\xi_j, \omega_j)'(\xi_j, \omega_j)|z] = \Omega(z_j),$$

with $\Omega(z_j)$ finite for almost every z_j.

The logic behind the estimation procedure is simple enough. Appendix 1 shows that given the data on the prices and the observed characteristics of the products, *any* choice of a triple consisting of an observed vector of positive market shares, say s, a proper distribution of consumer characteristics, say P, and the parameters of the model, say θ, implies a unique sequence of estimates for the two unobserved characteristics of our products, say $\{(\xi_j(\theta, s, P), \omega_j(\theta, s, P)\}_{j=1}^{J}$. We begin by assuming that we can actually calculate $\{(\xi_j(\theta, s_0, P_0), \omega_j(\theta, s_0, P_0))\}_{j=1}^{J}$ for alternative values of θ. In fact, we do not actually observe s_0 (though we do observe s_n), and for most of the models we consider we cannot actually compute the disturbances generated by P_0, but rather only from a (simulation) estimator of it. So our actual estimation procedure will be based on substituting estimates of s_0 and of P_0 into the algorithm we now develop.

Assuming we can compute $\{(\xi_j(\theta, s_0, P_0), \omega_j(\theta, s_0, P_0)\}$, then at $\theta = \theta_0$ our computation will reproduce the true values of the unobserved characteristics of the cars marketed. Consequently the conditional moment restrictions in (4.1) imply that any function of z must be uncorrelated with the vector $[\xi(\theta, s_0, P_0), \omega(\theta, s_0, P_0)]$ when that vector is evaluated at $\theta = \theta_0$. As in Hansen (1982), we can use this fact to generate a method of moments estimator of θ_0. That is, we can form the sample analog to a particular set of covariance restrictions and find that value of θ that sets this sample analog "as close as possible" to zero (see below).

The choice of moment restrictions is of interest for both statistical and economic reasons. First they determine the precision of our estimators. Also they serve to illustrate some of the differences between the determinants of prices and quantities in differentiated product models from their determinants in the more traditional homogeneous product model.

To be more precise we will need some additional notation. Let $T(z_j)$ be a 2 by 2 matrix of functions of z_j, and $H_j(z)$ be an L by 2 matrix of functions of z (the j index here indicates that the function may differ with the observation). Then if we define

$$(4.2)\ G(\theta) = E\left[H_j(z)T(z_j)\begin{pmatrix}\xi_j(\theta, s_0, P_0)\\ \omega_j(\theta, s_0, P_0)\end{pmatrix}|z\right],$$

(4.1) guarantees that $G(\theta_0) = 0$. Consequently if we form

$$(4.3)\ G_J(\theta; s_0, P_0) = J^{-1}\Sigma_{j=1}^{J} H_j(z)T(z_j)\begin{pmatrix}\xi_j(\theta, s_0, P_0)\\ \omega_j(\theta, s_0, P_0)\end{pmatrix},$$

and some mild regularity conditions are satisfied, $G_J(\theta_0)$ will converge in probability to zero and have a limit normal distribution. This suggests choosing, as our estimate of θ, the value which minimizes, up to a term of $o_p(1/\sqrt{J})$,

$$(4.4)\ \|G_J(\theta; s_0, P_0)\|\ ,$$

where for any vector y, $\|y\| = y'y$.

Appendix 1 provides conditions which insure that this estimator is root$-J$ consistent and asymptotically normal, with covariance matrix given by

$$(4.5)\ (\Gamma'\Gamma)^{-1}\Gamma'V\Gamma(\Gamma'\Gamma)^{-1},$$

where

$$\Gamma = \partial G(\theta_0)/\partial\theta, \text{ and } \ V = E_{(z)}H_j(z)T(z_j)\Omega(z_j)T(z_j)'H_j(z)'.$$

These conditions do not require us to be too specific about the nature of the dependence induced by the pricing equilibrium as J grows large.

Note that the precision of the implied estimate of θ will therefore depend on the choice of the matrices $T()$ and $H_j()$. These matrices play familiar roles in estimation. The matrix $T(\cdot)$ is introduced to standardize $[\xi(\theta_0), \omega(\theta_0)]$; so in the discussion that follows we will assume that

$$(4.6)\ T(z)'T(z) = \Omega(z)^{-1}.$$

$H_j()$, then, is a matrix of instruments for two standardized disturbances. These two disturbances are analogous to the demand and supply disturbances that appear in the traditional literature on the estimation of demand and supply curves in homogeneous product markets. They take a slightly more complicated form in our case both because the oligopolistic structure of the market implies that the cost side disturbance must be computed from the pricing equation rather than from the supply curve and because the discreteness in the choice set implies that we need a nonlinear transformation of the market shares to compute the demand side disturbance.

The issue of the appropriate choice of $H_j()$, then, is analogous to the issue of the choice of instruments in homogeneous product demand and supply systems. As in those systems, potential instruments for the demand equation include functions of the product's exogenous supply side characteristics (the $w_j's$), as well as functions of the exogenous demand side characteristics; and instruments for the pricing equation include functions of the product's exogenous demand side characteristics (the $x_j's$), as well as functions of the exogenous supply side characteristics. Note, however, that once we move into oligopolistic markets, then a firm's price depends not only on its own product's characteristics, but also, through the solution to the equilibrium in (3.5), on the characteristics of its competitors. Firms producing a good for which there are lots of close substitutes (products with similar characteristics) should face larger price elasticities (and hence choose lower price cost margins) than firms producing a good which does not. This suggests that the functions determining the instruments should depend on the attributes of other products, as well as the product's own attributes; or that $H_j(z)$ should depend on z_q for $q \neq j$, as well as on z_j. It does not, however, tell us the form of that dependence.

The recent literature on efficient estimation subject only to conditional moment restrictions provides guidance for choosing the instrument vector. Using an *i.i.d.* sampling scheme and other mild regularity conditions Chamberlain (1986) shows that the efficient set of instruments when we have only conditional moment restrictions is equal to the conditional expectation of the derivative of the moment condition with respect to the parameter vector (conditioning on the same set of variables which condition the moment restriction, and evaluated at θ_0). The analogous instruments for our case are

$$(4.7)\quad H_j(z) = E\left[\frac{\partial \xi_j(\theta_0, s_0, P_0)}{\partial \theta}, \frac{\partial \omega_j(\theta_0, s_0, P_0)}{\partial \theta} | z\right] T(z_j) \equiv D_j(z) T(z_j),$$

in which case the variance covariance matrix of the estimated parameter vector is

$$\{E_{(z)}[D_j(z)\Omega(z_j)^{-1} D_j(z)']\}^{-1}$$

The formula in (4.7) is very intuitive; it states that larger weights should be given to those observations which generate disturbances whose computed values are very sensitive to the choice of θ (at $\theta = \theta_0$). Unfortunately $D_j(z)$ is typically very difficult, if not impossible, to compute.

To compute $D_j(z)$ we would need to take the derivative of the integrand in (4.4) at different values of (ξ, ω), and then integrate out over the conditional distribution of these unobserved characteristics (conditioning on the observed exogenous variables). Since the unobserved characteristics, (ξ, ω), affect the values of the integrand through their impact on prices and market shares, to obtain the needed derivatives we would need to first compute equilibrium prices for the possible values of (ξ, ω). This in turn would require a rule to select among multiple equilibria when they exist. Moreover, to obtain the required integral of the derivative we would need additional assumptions on the conditional distribution of the unobserved components (or else a procedure for estimating it.) Though we have experimented with algorithms which compute approximations to $D_j(z)$ in this way, they have, thus far, proved to be too computationally demanding to be used repeatedly in our empirical work.[12]

Newey (1990) considers the special case where $T(z) = T$ (for all z), and shows that, again under mild regularity conditions, one can circumvent the problem of computing $D_j(z)$ by using a semiparametric estimator of it, and still generate an estimator whose limiting variance-covariance matrix is, $\{E_{(z)}[D_j(z)\Omega^{-1}D_j(z)']\}^{-1}$ (see also the related work on feasible GLS by Robinson,1987; and the literature cited in both of these articles). Newey(1990) also provides results from a Monte Carlo experiment which tends to show that this procedure works well when a polynomial series approximation to the efficient instrument vector is used.

Though polynomial approximations are easy to compute, there is a problem in using them to approximate functions that result from Nash games, such as ours. Typically the arguments in the function to be approximated from Nash games include the characteristics of all the competitors products. As a result an unrestricted polynomial series approximation of a given order will have a number of basis functions which grows polynomially in the number of other products in the market (our J). Formally then, our ability to use unrestricted polynomial approximations to optimal instruments for models with Nash equilibria differs in each of two cases. In the first, which is our case, the data consist of information on a large number of products in each of a few (in our case, only one) market. In this case J, the number of products, is also the limiting dimension of the problem. This implies that the dimension of the basis needed for the approximation grows polynomially

[12] To date we have concentrated on computing the needed derivatives at $(\xi, \omega) = (0, 0)$, and using these derivatives as instruments. One extension would use a "bootstrap" estimator of the distribution of (ξ, ω) to integrate out over the derivative evaluated at alternative possible of (ξ, ω). Note that all operational procedures for using the optimal instruments will involve using an approximation to their actual values, and hence require proofs that the approximation error that this induces does not affect the consistency or limit distribution of the estimator.

in sample size, and this, in turn, violates the regularity conditions required for the consistency of the first stage estimator of the efficient instruments. In the second case the data consist of information on a relatively small number of products in each of many markets. In this case the fact that we have repetitions over markets attenuates the dimensionality problem. In particular, if it is reasonable to assume that the relevant limits have the ratio of the number of markets to the number of products growing without bound, then the formal inconsistency in the econometric argument for using the polynomials disappears. Typically, however, even in the second case, once we include as arguments the factors that cause markets to differ from one another as well as the characteristics of all the products marketed, the dimension of the needed basis is still quite large relative to the number of observations.

As shown in Pakes (forthcoming) this dimensionality problem can be circumvented if the researcher is willing to assume that the chosen equilibrium is symmetric (more precisely exchangeable) in the state vectors of a product's competitors (that is, that we can permute the order in which one's competitors state vectors are listed in the data and not change the value of the given product's market share, or price). If the function we are trying to approximate is partially exchangeable (exchangeable in the state vectors of a product's competitors), it can be fit by a polynomial approximation which is also partially exchangeable. The space of exchangeable polynomials of a given order in J arguments is a subspace of the vector space consisting of all polynomials in J arguments of that order. Thus, if we use a partially exchangeable approximation we do not have to use a basis which spans the whole vector space of polynomials, but rather only a basis that spans the partially exchangeable subspace. Theorem 32 in Pakes (forthcoming) shows that the dimension of the basis for polynomials of a given order that are partially exchangeable is independent of the number of exchangeable arguments. The actual form of an easy to use basis for the exchangeable subspace is also provided, as is a table which provides an upper bound to the dimension of the basis (the bound is tight provided the number of competitors is greater than the order of the polynomial approximation).[13]

[13] We should note that there is very little experience with fitting exchangeable approximations to exchangeable functions, and what experience there is, is largely confined to problems with a different

Section (5.8) derives the exchangeable basis used in our problem. Let $f_j(z) \in \mathbf{R}^R$ provide the values of these basis functions for the j^{th} observation, $\otimes$ be the Kronecker product operator, and I_2 be an identity matrix of order two. To actually construct an estimator which uses an approximation to the efficient instruments it will be helpful to consider the special case in which $T(z) = T$ and the conditional expectation of the derivative matrix, $D_j(z)$, is a linear function of a finite dimensional basis. That is, $D_j(z)$ exactly equals $(f_j'(z) \otimes I_2)B$ for some matrix B. Let us consider an estimator which first projects the derivatives in (4.7) onto $(f_j'(z) \otimes I_2)$ and then uses the fitted values from this projection as the estimate of $D_j(z)$. It is possible to show that this estimator has the same limiting distribution as the generalized method of moments (or GMM) estimator (given in Hansen, 1982) which uses $\{[\xi(\theta), \omega(\theta)]' \otimes f_j(z)\}$ as moments and a consistent estimate of $E\{[\xi(\theta_0), \omega(\theta_0)]' \otimes f_j(z)\}\{[\xi(\theta_0), \omega(\theta_0)]' \otimes f_j(z)\}'$, as its weighting matrix. Since the method of moments estimator is easier to compute, we use it in the actual estimation subroutine.[14]

We now come back to the simplifying assumptions which we have maintained in the exposition of this section. First we need to account for the fact that we cannot actually compute the $G_J(\theta; s_0, P_0)$ in (4.4) needed to minimize the objective function in (4.5). There are really two separate problems here. The first is that we do not observe s_0 but just s_n, so all we can hope to calculate is $G_J(\theta, s_n, P)$. Second, for most of our models we will not be able to calculate $G_J(\theta; s, P_0)$ explicitly but will have to suffice with a simulation estimator of it. We show in the next section that this is equivalent to using $G_J(\theta; s, P_{ns})$ where P_{ns}

structure than ours. We are approximating a function at points which differ in only one component of the state vector (the component which determines which firm the instrument is being constructed for) whereas previous experience is with approximating functions at points which differ in the entire state vector (see Pakes, forthcoming, section V). Our empirical results tended to show that there is still some (incremental) predictive power in a low order exchangeable basis in our case, but the fits are not nearly as good as those reported in Pakes(forthcoming) for the case of approximating functions at points which differ in the entire state vector.

14 Additionally, if $D(z) \neq (f'(z) \otimes I)B$, but $T(z) = T$, then theorem 3.2 in Hansen (1982) insures that the GMM estimator is at least as efficient as the estimator which projects the derivatives on the finite basis to obtain its approximation to the optimal instruments. For the case $D(z) \neq (f'(z) \otimes I)$, and $T(z) \neq T$ we do not have an asymptotic efficiency ordering of the two estimators that holds over all data sets.

provides the empirical distribution of ns simulation draws from P_0[15] . Consequently, the objective function that our estimator θ actually minimizes is

$$(4.8) \quad \|G_J(\theta, s_n, P_{ns})\|.$$

The differentiability of the simulated market share function in θ and the boundedness of θ (see 5.13 below) imply that the simulated market shares converge to the theoretical market shares uniformly in θ over $\theta \in \Theta$ as ns grows large. This, together with the fact that $s_n - s_0 = O_p(1/\sqrt{n})$ and the differentiability of $\xi(\theta, s, P)$ in s, implies that provided $s_0 > 0$, $\sup_\theta \|G_J(\theta, s_n, P_{ns} - G_J(\theta, s_0, P_0)\|$ converges in probability to zero as both n and ns grow large. Recall that n in our sample is the number of households in the U.S. economy (a number on the order of 80 million.) Also, we have control over ns and, as we show below, have controlled it so that the simulation error has a negligible effect on our results. Consequently, we will ignore the effect of the substitution of $G_J(\theta, s_n, P_{ns})$ for $G_J(\theta, s_0, P_0)$ on the limiting properties of our estimators.

Finally we noted at the outset that the data we actually use is a panel data set which follows car models over all years they are marketed (and not a single cross section).[16] Moreover it is likely that the disturbances of a given model are more similar across years than are the disturbances of different models (so model-year combinations are not exchangeable). Though correlation in the disturbances of a given model marketed in different years does not alter the consistency or asymptotic normality of the parameter estimates from the algorithms described above, it does affect their variance-covariance matrix. As a result, we use estimators that treat the average of the moment restrictions of a given model over time as a single observation from an exchangeable population of models. That is, replacing product index j by indices for model m and year t, we define the sample moment condition associated with a single model as

$$g_m(\theta) \equiv \sum_t [f_{mt}(z)' \otimes I_2] \begin{bmatrix} \xi_{mt}(\theta) \\ \omega_{mt}(\theta) \end{bmatrix} .$$

---

15 Actually for increased efficiency we use an importance sampling simulator, so that $P(ns)$ becomes the empirical measure that results from ns draws from the importance sampling distribution, see the discussion in section (5.4).

16 Our definition of an automotive model is discussed in the data section below.

We then apply Hansen's GMM method to this sample moment condition. This then produces standard errors which allow for arbitrary correlation among models over time and arbitrary heteroskedasticity. Note, however, that while simple, this is not the most efficient method for dealing with correlation over time.

The computation can be further simplified. We note that the first-order conditions for this problem are linear in β and γ, for any given value of (α, σ). This allows us to "concentrate" β and γ out of the first-order conditions and perform a non-linear search only over the values of α and σ. To perform the required minimization, we use the Nelder-Mead (1965) non-derivative "simplex" method. We turn now to further details of our computational procedure.

5. Computation.

The method of moments estimation algorithm outlined in the last section requires computation of the demand and cost unobservables as a function of the model's parameter vector as well as our approximation to the optimal instruments. We now consider how to compute them, beginning with the unobservables, and then moving on to the instruments.

Throughout we focus on two special cases (and we present empirical results for both these cases in the next section). The first is the pure logit model, while the second adds interactions between consumer and product characteristics as in (2.7). The advantages of carrying along the logit model, despite the unreasonable substitution patterns that it implies, stem from its computational simplicity. This makes it easy to use the logit model to illustrate both the logic of the overall estimation procedure, and the likely importance of unobserved product characteristics.

Estimation of both models requires computation of the moment conditions, $G_J(\theta, s_n, P_{ns})$, evaluated at different values of the model's parameter vector, θ, and then a minimization routine which searches to find the value of θ which minimizes the objective function in (4.7). There are four steps needed to evaluate $G_J(\theta, s_n, P_{ns})$:

i) estimate (via simulation) the market shares implied by the model;

ii) solve for the vector of demand unobservables implied by the simulated market share function and the observed market shares, s_n;

iii) calculate the cost side unobservable, $\omega(\theta, s_n, P_{ns})$, as a function of the parameters of the model; and finally

iv) calculate the optimal instruments and interact them with the computed cost and demand side unobservables [as in (4.3)] to produce $G_J(\theta, s_n, P_{ns})$.

Both models we consider in detail are nested to the utility function,

$$(5.1)\ u_{ij} = \delta(x_j, p_j, \xi_j; \theta_1) + \mu(x_j, p_j, \nu_i; \theta_2) + \epsilon_{ij},$$

where the ϵ_{ij} are draws from independent extreme value distributions (independent over both i and j). Here $\delta_j = \delta(x_j, p_j, \xi_j; \theta_1)$ is a product-specific component that does not vary with consumer characteristics, while $\mu_{ij} = \mu(x_j, p_j, \nu_i; \theta_2)$ contains the interactions between product specific and consumer characteristics. We begin with the logit model.

5.1 The Logit Model.

Our first model will assume no interaction effects: *i.e.* $\mu_{ij} \equiv 0$. Given that we are assuming that ε_{ij} has the Weibull (or type I extreme value) distribution function, $\exp[-\exp(-\varepsilon)]$, the assumption that $\mu_{ij} \equiv 0$ gives us the traditional logit model for market shares. In addition to $\mu_{ij} \equiv 0$, we assume that the mean utility level is linear in product characteristics, or

$$(5.2)\ \delta_j = x_j\beta - \alpha p_j + \xi_j,$$

so that $u_{ij} = x_j\beta - \alpha p_j + \xi_j + \varepsilon_{ij}$. Recall that $u_{io} = \varepsilon_{io}$ (that is $\delta_0 = 0$), so the market-share functions are given by

$$(5.3)\ s_j(p, x, \xi, \theta, P_0) = e^{\delta}j \ /(1 + \textstyle\sum_{j=1}^{J} e^{\delta}j),$$

for $j = 0, 1, ..., J$ (see, for e.g. McFadden,1981).

The existence of the closed form in (5.3) helps explain the extensive use of the logit model. Also since (5.3) implies that

$$(5.4)\ \delta_j = \ln(s_j) - \ln(s_o),$$

the logit form's implied estimate of δ_j is $\ln(s_{nj}) - \ln(s_{no})$. That is, there are no computational problems in obtaining either the market shares or the inverse functions for this model (see i) and ii) above), and, as a result, we can obtain the demand-side unobservable analytically as

$$(5.5)\ \xi(s_n, p, x, \theta, P_0) = \ln(s_{nj}) - \ln(s_{no}) - x_j\beta + \alpha p_j.$$

We could interact the demand-side unobservables from (5.5) with instruments and apply a method of moments procedure to the resulting moment equations to estimate the demand-side parameters. This is the analog of the single equation instrumental variable estimator of the demand system in homogeneous product markets. For joint estimation of the demand and pricing equations we use the implied markups to compute the cost-side unobservables (see iii above).

To calculate the markups, we need the derivatives of market share with respect to price. In the logit case, these are

$$\frac{\partial s_j}{\partial p_r} = \begin{cases} -\alpha s_j(1 - s_j), & j = r; \\ \alpha s_j s_r, & j \neq r \ . \end{cases}$$

Note that in the logit model the markup depends only on the observed market shares and α; in fact, (3.4)-(3.6) imply that markups are inversely proportional to α. We can therefore write the logit markup as $b_j(s_n, \alpha) = \bar{b}_j(s_n)/\alpha$. If, temporarily, we maintain the assumption that marginal cost is linear in attributes (so that $c_j = \mathrm{w}_j\gamma + \omega_j$), then the cost-side unobservable is computed as

$$(5.6)\ \omega(s_n, p, x, \theta, P_0) = p - \mathrm{w}\gamma - \tfrac{1}{\alpha}\bar{b}(s_n).$$

Once again, the method of moments procedure could be applied to this equation alone (generating an instrumental variable estimator for the pricing equation), or we could try joint estimation of the moments generated from both (5.5) and (5.6).

Although we have noted that the logit model generates counter intuitive substitution patterns, note how easy it is to actually obtain estimates in the presence of endogenous prices in that model. In particular, estimation does not require either computation of nonanalytic integrals, or the computation of solutions to nonlinear systems of equations.

5.2 A Model with Interactions.

We now re-introduce a non-trivial interaction term $\mu_{ij} = \mu(x_j, p_j, \nu_i, \theta_2)$. For the reasons noted above we shall use the "Cobb-Douglas" specification in (2.7). However, the techniques developed here depend only on the (additional) assumptions that the consumer characteristics, the ν_i, have a known distribution (up to a parameter vector to be estimated), and that μ_{ij} does not depend on the unobservable ξ_j (although, as noted above, this last assumption can be relaxed at a computational cost).

It will be useful to obtain the market share function in two stages. First condition on the ν and integrate out over the extreme value deviates to obtain the conditional (on ν) market shares as

$$(5.7a) \quad f_j(\nu_i, \delta, p, x, \theta) = \frac{e^{\delta_j + \mu(x_j, p_j, \nu_i, \theta_2)}}{1 + \sum_{j=1}^{J} e^{\delta_j + \mu(x_j, p_j, \nu_i, \theta_2)}}.$$

Next integrate out over the distribution of ν to obtain the market shares conditional only on product characteristics as

$$(5.7\text{b}) \; s_j(p, x, \xi, \theta, P_0) = \int f_j(\nu_i, \delta(x, p, \xi), p, x, \theta) P_0(d\nu).$$

Note that (5.7a) has a closed form, while (5.7b) does not. Indeed we cannot compute (5.7b) exactly and will instead substitute a simulation estimator of its value into the estimation algorithm. Since the ϵ have been integrated out analytically, this allows us to confine our simulation procedure to the variance induced by the ν. In addition it produces an integrand for the simulation estimator which is a smooth function of all of its arguments. We come back to problem of efficiently simulating (5.7b) in the next subsection; for now we simply assume we have a good simulation estimator of it and label that estimator $s_j(p, x, \delta, P_{ns}; \theta)$.

Next we have to combine our estimates of the market share function with the observed market shares to solve for δ as a function of θ (see ii above). Once we add the interaction term, it becomes impossible to solve for δ analytically, so we will have to solve for it numerically each time we evaluate a different θ in the estimation algorithm. Recall that δ solves the non-linear system $s_n = s(\delta)$, or equivalently

$$\delta = \delta + \ln(s_n) - \ln(s(\delta)).$$

In appendix 2, we show that for any triple (s, θ, P), such that s is contained in the interior of the $J+1$ dimensional unit simplex, $\theta \in \Theta \subset \mathbf{R}^K$, and P is a proper distribution for ν, the operator $T_{(s,\theta,P)} : \mathbf{R}^J \rightarrow \mathbf{R}^J$ defined pointwise by

$$(5.8)\ T_{(s,\theta,P)}[\delta_j] = \delta_j + \ln(s_j) - \ln[s_j(p, x, \delta, P; \theta)],$$

is a contraction with modulus less than one. This suggests solving for δ recursively. That is we begin by evaluating the right-hand side of this equation at some initial guess for δ, obtain a new δ' as the output of this calculation, substitute δ' back into the right hand side of (5.8), and repeat this process until convergence.

Given $\delta_j(\theta, s, P)$, it is easy to solve for the demand-side unobservable as $\xi_j(\theta, s, P) = \delta_j(\theta, s, P) - x_j\beta$. Next we calculate the cost-side unobservable. To do so we need to solve for the markup and this requires the derivatives of the market share function with respect to price (see iii above). Given (5.7), those derivatives are easily shown to be,

$$(5.9\text{a})\ \partial s_j(p, x, \xi, \theta, P_0)/\partial p_j = \int f_j(\nu, \delta, x, p, \theta)(1 - f_j(\nu, \delta, x, p, \theta))[\partial(\delta_j + \mu_{ij})/\partial p_j]P_0(d\nu),$$

$$(5.9\text{b})\ \partial s_j(p, x, \xi, \theta, P_0)/\partial p_q = \int -f_j(\nu, \delta, x, p, \theta)f_q(\nu, \delta, x, p, \theta))[\partial(\delta_q + \mu_{iq})/\partial p_q]P_0(d\nu)\ .$$

5.3 Simulators for Market Shares.

As noted the integral in (5.7b) is difficult to calculate as K, the dimension of the consumer characteristics, grows much beyond two or three. As a result we form a simulation estimator of that integral and use it in the estimation algorithm. An easy way to do this is to replace the population density, $P_0(d\nu)$ in (5.7b), with the empirical distribution obtained from a set of ns pseudo-random draws from P_0, say, $(\nu_1, \ldots, \nu_{ns})$ and calculate

$$(5.10)\ s_j(p, x, \xi, \theta, P_{ns}) = \int f_j(\nu_i, \delta, p, x, \theta)P_{ns}(d\nu) \equiv \tfrac{1}{ns}\textstyle\sum_{i=1}^{ns} f_j(\nu_i, \delta, p, x, \theta).$$

The derivative in (5.10) is then estimated by $\partial s_j(p, x, \xi, \theta, P_{ns})/\ \partial p_j$, which has a similar simple analytic form.

Note that this simulation estimator has a smaller variance than the standard frequency simulator since much of the variance in consumer tastes, that owing to the "logit" errors ε_{ij}, has been analytically integrated out. On the other hand we should stress that the disturbance generated by the simulation process enters the moment condition defining our estimator in a non-linear fashion (so that consistency of our estimator requires ns/J to grow without bound as sample size grows large), and that the market shares we are trying to simulate are themselves quite small (see below for the data; the first order term in an expansion of the moment condition about the true market shares involves the disturbance from the simulation relative to that share). Consequently, we have looked for more precise simulators than that in (5.10) and done numerical comparisons of the accuracy of some of the alternatives. We turn to the results of that investigation now.

5.4 An Importance Sampling Simulator.

The importance sampling literature notes that we can often reduce the sampling variance of a simulation estimator of an integral by transforming both the integrand and the density we are drawing from in a way that reduces the variance of a simulation draw but leaves its expectation unchanged (see, for e.g., Rubinstein (1981), and the literature cited there). To see how to apply these techniques to our problem, consider any function $h(\cdot, \theta)$ which is strictly positive on the support of P_0. Assuming, for simplicity, that P_0 has a density with respect to Lebesgue measure, denoting that density by p_0, and suppressing some notation, the integral in (5.6b) can be rewritten as

$$(5.11)\quad s_j(\theta, P_0) = \int [\frac{f_j(\nu, \theta)}{h(\nu, \theta)} p_0(\nu) h(\nu, \theta)] d\nu \equiv \int f_{hj}(\nu, \theta) P_{hj}(d\nu, \theta) \equiv s_j(\theta, P_{hj}),$$

where

$$P_{hj}(d\nu, \theta) \equiv h(\nu, \theta) d\nu \text{ and } f_{hj}(\nu, \theta) \equiv [f_j(\nu, \theta) p_0(\nu)]/h(\nu, \theta).$$

This implies that any $h(\cdot)$ that is positive on the support of P_0 can be used to form an unbiased simulator of $s_j(\theta, P_0)$, say $s_j(\theta, P_{hj,ns})$, by rewriting the integral as in (5.11) and integrating out over the empirical distribution obtained from ns random draws from P_{hj}. It is well known that the choice for $h(\cdot)$ that generates a minimum variance estimator for that integral is

$$(5.12)\ P^*_{hj}(d\nu, \theta) = [f_j(\nu, \theta) p_0(\nu) d\nu] / \ s_j(\theta, P_0),$$

as, in this case, $s_j(\theta, P^*_{hj,ns})$ equals $s_j(\theta, P_0)$ exactly (no matter ns). Note that $P^*_{hj}(d\nu, \theta)$ places proportionately higher weight (relative to P_0) on draws of ν that result in larger values of the integrand. Intuitively, it tells us to over sample consumers whose characteristics would lead them to buy product j.

There are, however, well-known problems with this "optimal" simulator. Most importantly, its definition requires knowledge of the integral itself $[s_j(\theta, P_0)]$. Also, it depends on θ, while the limit properties of our estimator require the use of simulation draws which do not change as the minimization algorithm varies θ. Finally, note that the optimal importance sampling estimator changes with the share we are trying to simulate, while the contraction property that our inversion subroutine is based upon requires all the market shares to be simulated from the same simulator.

Though these problems make direct use of the simulator in (5.12) impossible, that formula does suggest how to build an importance sampling simulator with low variance. First note that though we do not know $P^*_{hj}(d\nu, \theta)$, we can obtain a consistent estimator of it, at least about $\theta = \theta_0$, by taking an initial consistent estimate of θ_0, say θ', calculating a good estimate of the share at θ', say $s_j(\theta', P_{nsi})$, and then drawing from $[f_j(\nu, \theta')p_0(\nu)d\nu]/\ s_j(\theta', P_{nsi})$. Note that the estimate $s_j(\theta', P_{nsi})$ is calculated only once, so nsi [the number of simulation draws for the initial step] can be quite large without imposing too much of a computational burden.

To implement this suggestion we need a way of drawing from $[f_j(\nu, \theta')p_0(\nu)d\nu]/s_j(\theta', P_{nsi})$. A simple *acceptance/rejection* procedure which accomplishes this is to draw ν from P_0 and "accept" it with probability $f_j(\nu, \theta')$. Bayes Rule implies that the accepted draws from this procedure have a density given by

$$P(d\nu/\text{accept}, \theta') = \frac{Pr(\text{accept}/\nu, \theta')P_0(d\nu)}{\int Pr(\text{accept}/\nu, \theta')P_0(d\nu)}$$
$$= \frac{f(\nu, \theta')P_0(d\nu)}{\int f(\nu, \theta')P_0(d\nu)} = \frac{f(\nu, \theta')P(d\nu)}{s_j(\theta', P_0)} \equiv P^*_{hj}(d\nu, \theta'),$$

as required.

Lastly since the contraction argument of appendix 2 requires that the same simulation draws be used to calculate each market share, we choose a market share on which to base the importance sampling simulator outlined above, and then use the simulator derived from it to simulate all shares. We focus on the share of households who purchase automobiles, that is, on $\bar{s}(\theta) \equiv [1 - s_o(\theta)] = \sum_{j=1}^{J} s_j$. This will lead us to over sample values of ν which are associated with high probabilities of purchasing an auto (as noted below, only about ten per cent of households buy a new car in any given year).

Thus we proceed as follows. We obtain an initial estimator of θ_0, say θ', using the simple smooth simulator in (5.10). Next we draw ν from P_0, accept it with probability $\bar{f}(\nu, \theta) \equiv \sum_{j=1}^{J} f_j(\nu, \theta)$, thereby generating an importance sampling simulator of the form

$$(5.13\text{a}) \quad P_{hs}^{*}(d\nu, \theta') = [\bar{f}(\nu, \theta') p_0(\nu) d\nu] / \bar{s}(\theta', P_0).$$

The vector of simulated market shares are then calculated as

$$(5.13b) \quad s_j[\theta, P_{hs}^{*}(\theta')_{ns}] = \sum_{i=1}^{ns} \frac{\bar{s}(\theta', P_0)}{\bar{f}(\nu_i, \theta')} f_j(\nu_i, \theta),$$

with the ν drawn from $P_{ns}^{h*}(\theta')$. Intuitively, then, we over sample (relative to P_0) the ν's which are more likely to lead to (some) auto being purchased and then weight the purchase probabilities, f_j, by $\bar{s}(\theta', P_0) / \bar{f}(\nu, \theta')$, the inverse of the sampling weights.

Section 5.6 provides a comparison of the performance of the alternative simulators. Before discussing these results, however, we need to explain how we incorporate exogenous information on the empirical distribution of consumer characteristics.

5.5 The Empirical Distribution of Income and the Final Form of the Simulator.

Recall that our interactive "Cobb-Douglas" model is written as

$$u_{ij} = \alpha ln(y_i - p_j) + x_j \bar{\beta} + \xi_j + \sum_k \sigma_k x_{jk} \nu_{ik} + \varepsilon_{ij}, \text{ for } \; j = 1, ..., J$$

while

$$u_{i0} = \alpha ln(y_i) + \xi_0 + \sigma_0 \nu_{i0} + \varepsilon_{i0}.$$

The vectors $(\nu_{i0}, ..., \nu_{iK})$, which determine the marginal utilities of characteristics, are assumed to be random draws from a normal with mean vector zero and an identity covariance matrix independent of the level of consumer's income (y_i). We treat the distribution of income differently because we have exogenous information on it from the March Current Population Survey (CPS) for each year of our panel.

It would be easy to take ns draws from the CPS for each year and simulate the market shares with the empirical distribution of these draws. However, this leads to a relatively imprecise simulator (in particular, it does a relatively bad job of estimating changes in the upper tail of the income distribution for the values of ns used in our empirical work). One alternative is to assume that the income distribution has a parametric form (we use log-normal), estimate the parameters of that distribution for each year from the income draws from the CPS (we denote the estimated mean by m_t and the estimated standard deviation by $\hat{\sigma}_{yt}$) and then simulate the market shares by drawing from a log normal with these parameters.

Using this procedure our utility model is written as

$$\text{(5.14a)}\ u_{ijt} = \alpha ln[\exp(m_t + \hat{\sigma}_y \nu_{iy}) - p_{jt}] + x_{jt}\bar{\beta} + \xi_{jt} + \textstyle\sum_k \sigma_k x_{jkt} \nu_{ik} + \varepsilon_{ijt},$$

$$\text{(5.14b)}\ u_{i0t} = \alpha ln[\exp(m_t + \hat{\sigma}_y \nu_{iy})] + \xi_{0t} + \sigma_0 \nu_{i0} + \varepsilon_{i0t},$$

where the vectors $(\nu_{iy}, \nu_{i0}, \ . \ . \ . \ , \nu_{iK})$ are random draws from a multivariate normal distribution with mean 0 and an identity covariance matrix. Note that now the only difference between our treatment of income and of the other consumer characteristics is that we take the parameters of the distribution of income from the CPS while we estimate the parameters of the distribution of the other consumer characteristics inside our algorithm. Other points to note are that we used $\hat{\sigma}_y$ instead of $\hat{\sigma}_{yt}$ (because changes in the latter appeared to be imprecisely estimated), and that we held the vector of characteristics $(\nu_{iy}, \nu_{i0}, \ . \ . \ . \ , \nu_{iK})$ fixed over the time period of the panel.

5.7 Estimates of The Variance Induced by the Simulation Process.

Since the variances of the alternative simulation estimators of market shares, and their implications for the induced simulation errors in the mean utility levels (in δ), depend on the precise values of the parameters of the model, we calculated these variances at the estimated parameter values of our "base case" model. These parameter estimates are reported below in Table 6 (and discussed at length in section 6).

We contrast the performance of the importance sampling simulator in (5.13), with the performances of the smooth simulator in (5.10) and the naive "frequency" simulator. The latter is defined by taking draws from the distribution of both ν_i and ε_{ij} and then calculating the proportion of individuals who purchase a given car. Table 1 summarizes our results. The variance of the frequency estimator with mean s_j is $s_j(1 - s_j)$, while the variance of the two more complicated methods are calculated from 1000 draws from either P_{ns} or $P^*_{hs,ns}$ (as defined in 5.13) as appropriate.[17] Note that we are not primarily concerned with the impact of the simulation on the estimates s_j, but rather with its impact on the δ_j, as these are the variables that enter the estimation algorithm. To measure the impact of the simulation process on the estimates of δ we calculated δ 100 times, once from each of 100 independent samples of 200 draws on the consumer taste characteristics. We then calculated the empirical standard deviations of these 100 vectors.

The numbers in the table reflect the simulation error in the s_j's and δ_j's for the 137 products marketed in 1990. Row's 1 and 2 of the table provide the mean (across products) of the calculated standard deviations of s_j and δ_j respectively. The first row of Table 1 indicates that the mean (across products) of the standard deviation of the simulation error in market shares calculated from the frequency simulator is about 7 times as large as that from the simple smooth simulator and is more than 25 times larger than the standard deviation of the importance sampling simulator. The numbers in this row are for the standard deviation from one random draw. To obtain the standard deviation associated with the empirical estimates one must divide them by the square root of the number of draws (200) used to obtain those results. The result for the importance sampling simulator is a standard deviation on the order of 0.0001, which seems quite small. Indeed, a simple comparison of the variances of the frequency and importance sampling draws indicates that more than 100,000 "frequency" draws would be required to reduce the mean standard deviation of the simulation error in the market shares to the level achieved by 200 draws

17 Note that the optimal importance sampling distribution varies across years, whereas we want a set of simulation draws which are constant across years. In the empirical work reported below, we take 10 draws from the "optimal" distribution for each of 20 years, for a total of 200 draws. To calculate the standard deviations of the market shares reported in column 3 of Table 1, we took 50 draws from the "optimal" distribution for a total of 1060 draws. In each case, we use the full set of draws to calculate the market shares in each year.

in the importance sampling method. Remember, however, that the mean value (across products in 1990) of observed market shares is about 0.0007 (with a standard deviation of about 0.0008). So the 0.0001 number may not be so small relative to the market shares. Whether it is small enough to be ignored depends on its effect on δ. This effect must take account of both the derivative of each δ with respect to the entire vector of market shares, and the (negative) covariance among market shares induced by the simulation process.

The second row of the table gives the standard deviation of the simulation error in δ_j for the simple smooth and importance sampling estimators. (Because the frequency simulator is discontinuous, δ is not well-defined by the "frequency" market shares, and so we do not calculate the variance in δ with respect to that simulator.)

To answer the question of whether the simulation process induces an error with a significant variance in δ, we calculated the "signal to noise" ratio in the calculated δ_j's by taking the ratio of the variance in "true" δ_j's across products to the mean variance of the simulation error in the δ_j. The mean of the calculated $\delta_j's$ is -1.35 with a standard deviation of 2.05. The signal to noise ratio, with 200 simulation draws, is thus slightly more than 45 (or, the noise to signal ratio is about 0.022.) That is, with 200 simulation draws from our important sampling simulator, only about two percent of the total variance in the δ_j's is due to the simulation error (the rest being due to observed and unobserved product attributes). We think this is small enough to ignore.

5.8 Instruments.

Just as in homogeneous product demand and supply systems, in order to provide consistent estimates of the demand and pricing equations from our differentiated product model we require instruments for prices. As noted in section 4, the results in Chamberlain (1986) indicate that the optimal instruments for our problem involve the conditional expectation of the derivative of the disturbance vector with respect to the parameters of interest, conditional on the set of exogenous variables for the model. Because this conditional expectation is difficult to compute, we approximate it by a polynomial in the relevant variables. We assume that the Nash equilibrium in prices is symmetric, or more precisely, exchangeable in ways which we will detail below. We then restrict the approximation to satisfy these same properties and show how these restrictions make it possible to use polynomial approximations to the efficient instrument function in models with Nash equilibria.

To obtain an unrestricted q^{th} order polynomial approximation to a function, we form a basis that is rich enough to span the space of all polynomials of order q, and then look for the element of the vector space spanned by this basis that "best fits" the function we are after (throughout we will, for simplicity, use Euclidean distance). The advantage of restricting the approximation to be within a subspace of the original vector space, in our case the subspace of exchangeable polynomials, is that the dimension of the basis needed to span the subspace can be smaller than the dimension of the basis needed to span the original space. The characteristics of the restricted basis will depend on the nature of the restrictions that the functions being approximated satisfy. We now specify those restrictions.

Let $z_f = [z_{1f}, , ..., z_{Jf}]$, with $z_{jf} \in \mathbf{R}^{K+1}$, be the vector of observed characteristics of the J products marketed by firm "f". Note that this notation assumes that all firms market the same number of products; below we generalize to allow for different numbers of products offered by different firms. Let $\psi(\cdot)$ be the map from observed product characteristics to the (optimal) instruments defined in (4.7). It can be shown that a symmetric Nash equilibrium implies that this function is exchangeable in:

a) the characteristic vectors of the products marketed by a firm's competitors, *i.e.* for any $z = (z_1, ..., z_F)$

$$(5.15a)\ \psi(z_1; z_2, ... z_F) = \psi(z_1; z_{\pi(2)}, ..., z_{\pi(F)}) \ ,$$

and any permutation, $[\pi(2), ..., \pi(F)]$ of $[2, ..., F]$;

b) for a given competitor, it is exchangeable in the characteristic vectors of that competitor's products, or

$$(5.15b)\ \psi[z_1; ..., (z_{1f}, , ..., z_{Jf}), ..., z_F] = \psi[z_1; ..., (z_{\pi(1)f}, , ..., z_{\pi(J)f}), ..., z_F]$$

for any permutation $[\pi(1), ..., \pi(J)]$ of $[1, ..., J]$, and any $f \in [2, ..., F]$; and

c) for a given product, it is exchangeable in the characteristics of the other products of the firm marketing it, or

$$(5.15c)\ \psi[(z_{11}, z_{21}, ..., z_{J1}); z_2, ..., z_F) = p[(z_{11}, z_{\pi(2)1}, ..., z_{\pi(J)1}); z_2, ..., z_F),$$

for any permutation $[\pi(2), ..., \pi(J)]$ of $[2, ..., J]$.

Let $\alpha(h_1; h_2, ..., h_F)$ be the coefficient associated with the polynomial basis function which has powers given by the vector h_f for the coefficients of z_f in the polynomial approximation of $\psi(\cdot)$. Then, as shown in Pakes (forthcoming, lemma 33.2) if the approximating function is to abide by the restrictions of (5.15) these coefficients will satisfy

$$(5.16a)\ \alpha(h_1; h_2, ..., h_F) = \alpha(h_1; h_{\pi(2)}, ..., h_{\pi(F)})\ ,$$

for any permutation, $[\pi(2), ..., \pi(F)]$ of $[2, ..., F]$;

$$(5.16b)\ \alpha[h_1; ..., (h_{1f}, , ..., h_{Jf}), ..., h_F] = \alpha[h_1; .., (h_{\pi(1)f}, , ..., h_{\pi(J)f}), ..., h_F]$$

for any permutation $[\pi(1), ..., \pi(J)]$ of $[1, ..., J]$, and any $f \in [2, ..., F]$; and

$$(5.16c)\ \alpha[(h_{11}, h_{21}, ..., h_{J1}); h_2, ..., h_F) = \alpha[(h_{11}, h_{\pi(2)1}, ..., h_{\pi(J)1}); h_2, ..., h_F),$$

for any permutation $[\pi(2), ..., \pi(J)]$ of $[2, ..., J(1)]$.

One way of seeing the impact of the restrictions in (5.16) on the dimension of the basis is by providing the first order basis functions from a polynomial which satisfies them. These can be obtained by summing the first order basis functions from a standard polynomial basis that, from (5.16), are restricted to have the same coefficients. Let the coefficients of the unrestricted first order basis functions be denoted by $\{\alpha_{kjf};$ for $k = 1, ..., K+1$, $j = 1, ..., J$, and $f = 1, ..., F\}$. Then the restrictions in (5.16) imply that

$$\alpha_{kjf} = \alpha_{kj\cdot} \text{ for } f \in [2, ..., F], \text{ and all } (k, j)$$

$$\alpha_{kjf} = \alpha_{k\cdot\cdot} \text{ for } j \in [2, ..., J(f)], f \in [2, ..., F], \text{ and all } k, \text{ and}$$

$$\alpha_{kj1} = \alpha_{k\cdot 1} \text{ for all } k.$$

This in turn implies that the first order basis functions associated with characteristic "k" are

$$(5.17)\ z_{k11},\ \Sigma_{j=1}^{J} z_{kj1},\ \text{and}\ \Sigma_{f=2}^{F}\Sigma_{j=1}^{J} z_{kjf}.$$

Thus the dimension of the first order terms in the restricted basis is $3(K+1)$. In contrast, the dimension of the first order terms in the unrestricted basis is $FJ(K+1)$. Note that the number of first order basis functions required in the restricted basis is independent of both the number of firms and the number of products marketed by each firm. The content of theorem 33.1 in Pakes (forthcoming) is that this will be true for the number of basis functions of any order. Since our empirical results use only the first order terms from the polynomial approximations, and the actual number of basis functions required for computing a restricted basis of a order $q > 1$ requires some rather detailed combinatorics, we suffice with (5.17) here.

We now return to the fact that different firms market different numbers of products. Probably the easiest way to deal with this possibility is to hold J fixed at its maximum value, and introduce an indicator function for each of the J potential products of each firm which equals one only if that product is marketed (with the understanding that the other characteristics of products not marketed are set to zero). If we then go through the formalities of the restricted polynomial approximation given above the indicator function (which is now one of the $K+1$ characteristics) will generate the following three basis functions (see 5.17); one (the constant term), the number of products marketed by the firm, and the sum of the number of products marketed by its competitors. We should note, however, that though we use this simple way of adjusting the polynomial approximation for the fact that different firms market different numbers of products, it need not always be appropriate. In particular, it is essentially relying on a smooth approximation to the effect of a discrete valued variable (the number of products marketed by the firm), and this might not provide a very good approximation, especially if there are only a small number of products marketed by each firm.[18]

[18] Note also the implicit assumption that the number of products marketed by the firm is exogenous; that is functions of it are assumed uncorrelated with the demand and supply side disturbances for the firm's products. To control for the approximation problems induced by the fact that different firms market

different numbers of products without using a smooth approximation to this discrete variable, we can go back to the original problem and only restrict the approximation to the instrument function to be exchangeable in the state vectors of the competitors that market the same number of products. This would generate different restricted polynomial approximations for groups of firms defined by the number of products they market. If the number of products were large enough one might expect those coefficients to be relatively "smooth" functions of the number of products a firm markets. By generating a state variable that equals the number of products marketed by the firm, what the technique suggested above is doing is allowing those coefficients to be approximated by a polynomial in the number of products marketed.

Automobile Prices in Market Equilibrium: Part II.

Steven Berry
Yale University
National Bureau of Economic Research
James Levinsohn
University of Michigan
National Bureau of Economic Research
and
Ariel Pakes
Yale University
National Bureau of Economic Research

6. Data and Results.

6.1 The Data.

We use data on product characteristics obtained from annual issues of the *Automotive News Market Data Book.*[1] Product characteristics for which we have data include the number of cylinders, number of doors, weight, engine displacement, horsepower, length, width, wheelbase, EPA miles per gallon rating (MPG), and dummy variables for whether the car has front wheel drive, automatic transmission, power steering, and air conditioning as standard equipment.

The price variable is the list retail price (in $1000's) for the base model. This is clearly not ideal; we would prefer transactions prices, but these are not easy to find. All prices are in 1983 dollars. (We used the Consumer Price Index to deflate). The sales variable corresponds to U.S. sales (in 1000's) by name plate. The product characteristics correspond to the characteristics of the base model for the given name plate.

We would like to thank Don Andrews, Tim Bresnahan, Gary Chamberlin, Zvi Griliches, Jerry Hausman, Whitney Newey, and Frank Wolak for helpful comments. While working on this paper, Berry was an Olin Fellow at the NBER and Levinsohn was a National Fellow at the Hoover Institution, Stanford University. Each thanks their respective hosts. We gratefully acknowledge funding from National Science Foundation Grants SES-8821722 (to Richard Erikson and Ariel Pakes) and SES-9122672.

[1] The data set combines data collected by us with a similar data set graciously made available to us by Ernie Berndt of MIT.

The data set includes this information on (essentially) all models marketed during the 20 year period beginning in 1971 and ending in 1990 (the only models excluded are "exotic" models with extremely small market shares, such as the Ferrari and the Rolls Royce). Since models both appear and exit over this period, this gives us an unbalanced panel. Treating a model/year as an observation, the total sample size is 2217. Throughout we shall assume that two observations in adjacent years represent the same model if (a) they have the same name; and (b) their horsepower, width, length, or wheelbase do not change by more than ten percent. With these definitions the 2217 model/years represent 997 distinct models (as noted in section 4, different models are assumed to have unobservables whose conditional distributions are independent of one another, but the unobservables for different years of the same model are allowed to be freely correlated).

Aside from these product characteristics, we obtain additional data from a variety of sources. Because we thought that the cost of driving may matter to consumers (as opposed to just the MPG rating), we gathered data on the price of gasoline (the real price of unleaded gasoline as reported by the U.S. Department of Commerce in *Business Statistics*, 1961-1988). One of our product characteristics is then miles per dollar (MP\$), calculated as MPG divided by price per gallon. Also our measure of market size (M) was the number of households in the U.S. and this was taken for each year from the *Statistical Abstract of the U.S.*, while, as noted in the computation section, the parameters of the distribution of household income were estimated from the annual March Current Population Surveys. We also obtained Consumer Reports reliability rating for each model. This variable is a relative index that ranges from 1 (much less than average reliability) to 5 (much better than average reliability.)

The multi-product pricing problem requires us to distinguish which firms produce which models. We assume that different branches of the same parent company comprise a single firm. For example, Buick, Oldsmobile, Cadillac, Chevrolet, and Pontiac are all part of one firm, General Motors. This follows Bresnahan (1981) and Feenstra and Levinsohn (1991). For some results, we also assign a country of origin to each model, which is simply the country associated with the producing firm.[2]

[2] For example, we treat Hondas as Japanese and VW's as German, although, by the end of our

Tables 2 and 3 provide some summary descriptive statistics of variables that are used in the specifications we discuss below. These variables include quantity (in units of 1000), price (in $1000 units), dummies for where the firm that produced the car is headquartered, the ratio of horsepower to weight (in HP per 10 lbs.), a dummy for whether air conditioning is standard (1 if standard, 0 otherwise), the number of ten mile increments one could drive for $1 worth of gasoline (MP$), tens of miles per gallon (MPG), and size (measured as length times width). Table 2 gives sales-weighted means. Several interesting trends are evident. The number of products available generally rises from a low of 72 in 1974 to its high of 150 in 1988. Sales per model, on the other hand trend downward (though here there is some movement about the trend). In real terms, the sales-weighted average list price of autos has risen almost 50 percent during the 1980s after having remained about constant during the 1970s. On the other hand, the characteristics of the cars marketed are also changing (so the cost of a car with a given vector of characteristics need not be increasing). The ratio of horsepower to weight fell in the early 1970s and has since trended upward. Most of the changes in this ratio are attributable to changes in weight as horsepower has remained remarkably constant. It appears that prior to the first oil price shock, cars were becoming heavier and after the mid-1970s, cars became lighter. Along with the change in the ratio of horsepower to weight, cars have also become more fuel cost-efficient. In 1971, the average new car drove 18.50 miles on a (1983) dollar of gasoline, while by 1990 that figure was 28.52 miles. Also, while no cars had air conditioning as standard equipment at the start of the sample, 30.8 percent had it by the end. This is indicative of a general trend toward more extensive standard equipment. The market share of domestic cars has fallen from a 1973 high of 93.2 percent to a 1990 low of 68.2 percent. European market share has been fairly constant since the demise of the popular VW Beetle in the mid-1970s hovering around 4 to 5 percent. The Japanese market share has risen from a low of 4.0 percent in 1973 to a high of 27.6 percent in 1990. An automobile's size, given by its length times width trends generally downward with this measure falling about 17 percent over the sample.

---

sample, some of each were produced in the U.S.

Table 3 associates some names with the numbers. This table provides an indication of the range of the continuous product attributes by presenting the quartiles of their distribution. The least expensive car in the sample is the 1990 Yugo at $3393 (1983 dollars) while the top-of-the-line Porsche 911 Turbo Cabriolet costs $68,597. The 1989 Geo Metro has the highest MPG and MP$ while the 1974 Cadillac Eldorado has the lowest. The ratio of horsepower to weight varies tremendously from 0.170 for the (questionably named) 1985 Plymouth Gran Fury to .948 for the Porsche 911 Turbo. The smallest car in the sample was the 1973 Honda Civic. Tables 2 and 3 illustrate the tremendous variance in the sample both over time and across characteristics.

6.2 *Some Results.*

We will report three basic sets of results together with some auxiliary calculations. These are a simple logit specification, an instrumental variables logit specification, and the Cobb-Douglas specification in (5.14) above (Cobb-Douglas in the income available for the purchases of other goods and an index of the value of the attributes of the product, with random coefficients on the attribute vector, see 5.14 above). For simplicity, we will refer to the first as logit, the second as IV logit, and the third as BLP. The logit results provide an easy to compute reference point and will be discussed first. The IV logit maintains the restrictive functional form of the logit (and hence must generate the restrictive substitution patterns that this form implies), but allows for unobserved product attributes that are correlated with price, and therefore corrects for the simultaneity problem that this correlation induces. The BLP results allow both for unobserved product characteristics and a more flexible set of substitution patterns. Results from each specification will be discussed in turn.

6.3. *The Logit and the IV Logit.*

The first set of results are based on a standard logit specification for the utility function. They are obtained from an ordinary least squares regression of $\ln(s_{nj}) - \ln(s_{n0})$ on product characteristics and price (see 5.5).

The choice of which attributes to include in the utility function is, of course, *ad hoc*. For the BLP specification, computational constraints dictate a parsimonious list. Since we wish to compare results across different specifications, we adopt a short list of included attributes in the logit specifications also. Included characteristics are the ratio of horsepower to weight (HPWT), a dummy for whether air conditioning is standard, miles per dollar (MP$), size, and a constant. Horsepower over weight and MP$ are obvious measures of power and fuel efficiency, while air conditioning proxies for a measure of luxury. Size is intended as a measure of both itself and safety. Other measures of size such as interior room are not available for much of the sample period while government crash test results are only available for a small subsample of the data. Though there are surely solid arguments for including excluded attributes, their force is somewhat diminished by our explicit treatment of product attributes unobserved by the econometrician but known to the market participants. Still, we investigate how robust results are to the choice of included attributes in sensitivity analyses that are presented below.

In the first column of Table 4, we report the results of OLS applied to the logit utility specification. Most coefficients are of the expected sign, although the (imprecisely estimated) negative coefficients on air conditioning and size are anomalies, as one would expect these attributes to yield positive marginal utility. On the other hand these estimates have a distinctly implausible set of implications on own price elasticities. The estimated coefficient on price in Table 4 implies that 1494 of the 2217 models have inelastic demands. This is inconsistent with profit maximizing price choices. Moreover this is not simply a problem generated by an imprecise estimate of the price coefficient. Adding and subtracting two times the estimate of the standard deviation of the price coefficient to its value and recalculating the price elasticities still leaves 1429 and 1617 inelastic demands respectively.

In the second column of Table 4, we re-estimate the logit utility specification, this time allowing for unobservable product attributes that are known to the market participants (and hence can be used to set prices), but not to the econometrician. To account for the possible correlation between the price variable and the unobserved characteristics, we use an instrumental variable estimation technique. The instruments used are the nonparametric (or exchangeable) instruments discussed in section 5.8. They are functions of both the product's own attributes and the attributes of other products.

The use of instruments generates substantial changes in several of the parameter estimates. All characteristics now enter utility positively and all but MP$ are statistically significant. Moreover, just as the simultaneity story predicts, the coefficient on price increases in absolute value. (Indeed it more than doubles). Our interpretation of this finding is simply that products with higher unmeasured quality components sell at higher prices. Note that now only 22 products have inelastic demands – a significant improvement from the OLS results. Seven to 101 demands are estimated to be inelastic when we evaluate elasticities at plus and minus two standard deviations of the parameter estimate.

These results seem to indicate that the correction for the endogeneity of prices matters. One can also see the scope for unobservable product characteristics by examining the fit of the logit demand equation. The simple logit specification gives an R^2 of 0.387. This implies that 61 percent of the variance in mean utility levels is due to the unobserved product characteristics.

Though the use of IV estimation techniques in conjunction with the logit does seem to generate more plausible own price elasticities of demand, as noted in section 2, the fact that the functional form specification in the logit is separable in product and consumer characteristics implies that neither the IV nor the simple logit estimates can possibly generate plausible cross price elasticities, or for that matter differences in markups across products. This point is illustrated in table 5.

Table 5 provides a number of the semi-elasticities implied by the IV estimates reported in Table 4.[3] Each semi-elasticity gives the percentage change in market share of the row car associated with a $1000 increase in the price of the column car. The table is most easily interpreted by reading down the columns. It is immediately obvious that an increase in the price of any given model has the same semi-elasticity with respect to every other model. For example, a $1000 price increase of the top-of-the-line BMW (735i) increases the market share of the Ford Escort by the same percentage amount as it increases the market share of the Lexus LS400. Alternatively, under the logit specification when the price of an auto increases, the consumers that substitute out of that auto are more likely to substitute towards products with high market shares, regardless of the characteristics

[3] We only report results in table 4 and in subsequent tables for only a handful of models, which were chosen to represent the range of models in the sample. We list the same models in each case (tables 5, 7, 8, 9, 10, and 11.) Full results are available upon request.

of the product whose price increased. Thus, as one reads across rows the values of the cross-price terms vary proportionately with observed market share; for every product, an increase in the price of the Honda Accord (the top selling car in the table) will cause the largest percentage change in sales.

Finally, note that the own-price semi-elasticities are almost all identical. This too follows from the functional form of utility, as the own price semi-elasticity is just given by $\alpha(1-s_j)$. As the market shares (out of about 80 million households) never exceed 0.0068, own-price elasticities are necessarily very similar. The model is simply unable to generate reasonable substitution patterns, and no data could change this result.

In an oligopoly context unreasonable patterns of demand elasticities translate into unreasonable patterns of markups. Based on the IV logit estimates reported in Table 4, we find that all models have about the same mark-up, ranging from \$4630 for the BMW to \$4805 for the Chevy Cavalier. Markups are related to the model's market share (which, as noted, are about equal in absolute terms for all products) and how many products are made by the same parent firm. GM produces the most models and its markups are highest, while BMW produces the fewest models and their markups are, quite counter-intuitively, the lowest.

The IV logit estimates that these tables were based upon were obtained by estimating the demand system in isolation. As noted in section 4, if we are willing to specify a pricing game, it is possible to derive more efficient estimates of the parameters of the demand system by jointly estimating the demand and the pricing equations. The markup terms that are needed for this exercise are determined by the own- and cross-price elasticities obtained from the demand system. We thought that the substitution patterns that emanated from the logit demand system were too unbelievable to make joint estimation of the demand and pricing equations interesting in this context.

On the other hand, there is at least one familiar base case pricing equation, the marginal cost pricing equation, which does have its own inherent logic, and therefore ought probably to be compared to the estimates we later obtain from the BLP specification. The marginal cost pricing equation is obtained by setting the markup term in our pricing equation (equation 3.6) to zero, and regressing log price on w (the characteristics which shift the cost surface). This just amounts to setting price equal to marginal cost and estimating the marginal cost equation in 3.1. Note that the pricing equation obtained in this fashion is just a familiar hedonic pricing equation (see Griliches,1971).

The third column of Table 4 presents the results from OLS estimation of equation 3.1 with $\ln(c_j)$ set equal to $\ln(p_j)$. In Table 4 (and in subsequent cost-side results), included cost shifters (w_j) are the same attributes that appear in utility with three modifications. First, miles per gallon replaces miles per dollar, as the production cost of fuel efficient vehicles presumably does not change with the retail price of gasoline (at least in the short-run). Second, we include a trend term to capture technical change and other trending influences (*e.g.* government regulation) on real marginal cost. Third, we use the log of continuous attributes, not their level, in the cost function. Thus the cost function parameters have the interpretation of elasticities of marginal cost with respect to associated product characteristics.

Note that the cost function adopted here is both simple and restrictive. In particular, it implies a constant elasticity of marginal cost with respect to all attributes and does not permit marginal cost to vary with output. Though our robustness tests provide some results with more flexible cost functions (see Table 10 below), we were hesitant about using a more complex specification for the cost surface without having more direct information on costs (see the discussion of extensions in section 7 below).

As is typical in hedonic pricing regressions, each of the coefficients on characteristics (except MPD) is estimated to be positive and all are significantly different from zero. (We come back to comment on the MPD coefficient below.) For example, a 10% increase in the ratio of horsepower to weight is associated with a 5.2% increase in prices (and, in this context, in marginal costs). Also familiar from hedonic results is the fact that the R^2 from this regression is fairly high (at 0.66); simple functions of observable characteristics seem to be much better able to explain differences in the log of prices, than they are able to explain differences in the mean utility levels that rationalize the logit demand structure.

We turn now to results from our full model.

6.3 *Results from the Full Model.*

The demand system for the full model is derived from the utility function in (5.14). Recall that this is Cobb-Douglas in the income available for the purchases of other goods and in an index of the value of the product's characteristics. Differences in choices among individuals are generated by individual-specific differences in income and in the marginal utilities of observable characteristics, together with an *i.i.d.* product-specific extreme value deviate. The joint distribution of the log of income and the marginal utilities is assumed to be normal and independent across characteristics with the parameters of the income distribution estimated from the CPS, and the parameters of the marginal utility distribution estimated inside our algorithm. The pricing equation (given in 3.6) is derived from the assumptions that equilibrium is Nash in prices and that the marginal cost is log linear in attributes (and constant in quantities).

The attributes which enter the utility function (the x-vector) for our base case scenario are the same as in Table 4: a constant, HP to weight, air conditioning, MP\$, size, and a constant. Now, the marginal utility of each attribute varies across consumers so that we will be estimating a mean and a variance for each of them. In this context, we remind the reader that a positive variance of the random coefficient on the constant term implies that the distribution of the outside good has more idiosyncratic variance than that of the extreme value deviates generating idiosyncratic variance for the inside alternatives.

We construct the instrument function $H_j(z)$ of section 4 from the linear terms in the exchangeable basis (see section 5.8) generated by the exogenous supply and demand variables. We construct separate vectors of demand instruments, which are interacted with the demand unobservable, $\xi(\theta)$), and cost instruments, which are interacted with $\omega(\theta)$. The demand instruments consist of the x_j vector itself, the sum of the elements of x_j across products produced by the firm that produces j and the sum of the elements of x_j across the products of other firms. Since there are 5 demand-side variables, this gives 15 demand side instruments. Any variable that affects cost but not demand could be added to this list of instruments. However, we find that the only such variable in our specification, miles per gallon, is nearly collinear with the first 15 variables and so we do not include it. The cost side instruments are constructed in the same way from w_j. There are 6 elements of w_j, giving at least 18 elements cost side instruments. We were able to add the excluded demand variable, miles per dollar, to this list without causing a problem with near collinearity. Therefore, there are 19 cost side instruments, giving a total of 34 $(15 + 19)$ sample moment restrictions.

The results from jointly estimating the demand and pricing equations from our specification are provided in Table 6. The reported standard errors have been corrected for serial correlation of unobserved characteristics within models across years but not for any correlation across models.[4] The first and second panels of the table provide the estimates of the means and standard deviations of the taste distribution of each attribute, respectively. The third panel provides the estimate of the coefficient of $\ln(y - p)$, and the last panel provides the estimates of the parameters of the cost functions.

We begin with a discussion of the cost-side parameters. The cost function variables are the same as in the third column of Table 4: a constant, ln(HP/weight), AC, ln(MPG), ln(size), and a trend. The coefficients on ln(HP/Weight), Air, and the constant are positive and significantly different from zero. The term on trend is also positive and significant. The coefficient on ln(size) is not significantly different from zero. The coefficient on MPG is negative and significant, just as it is in the marginal cost pricing equation in table 4,

Indeed, recall that our pricing equation is essentially an instrumental variable regression of $ln[p - b(p,x,\xi;\theta)]$ on the cost side characteristics, where $b(p,x,\xi;\theta)$ is the markup (see 3.6). Since $ln[p - b(p,x,\xi;\theta)] \cong ln(p) - b(p,x,\xi;\theta)/p$, if our model is correct, the marginal cost pricing, or "hedonic," regression should, by the traditional omitted variable formula, produce coefficients which are approximately the sum of the effect of the characteristic on marginal cost and the coefficient obtained from the auxiliary instrumental variable regression of the percentage markup on the characteristics. Comparing the cost side parameters in table 6 with the hedonic regression in table 4, we find that the only two coefficients which seem to differ a great deal between tables are the constant term and the

[4] We should note here that we have also estimated the demand side of our specification separately, and that we have run specifications that allowed for firm specific dummy variables on both the demand and cost side. Since there are 22 firms in our data set this latter specification generates 66 additional parameters (a mean and variance for each firm on the demand side, and one cost elasticity on the supply side). Neither of these changes generated point estimates that were much different from the point estimates in table 6, but both generated estimated standard errors which were much larger than those in that table.

coefficient on size. The fall in these two coefficients should just be telling us that there is a positive average paercentage markup, and that this markup tends to increase in size.

The coefficients on MPG and size are probably a result of our constant returns to scale assumption. Note that sales (or production) does not enter the cost function. This is due to data limitations. Almost all domestic production is also sold in the U.S., hence domestic sales is an excellent proxy for production. The same is not true for foreign production, and we do not have data on model level production for foreign automobiles. The negative coefficient on MPG may result because the best selling cars are also those that have high MPG. By imposing constant returns to scale, we may force these cars to have a smaller marginal cost than they actually do. Due to the positive correlation between both MPG and size and sales, conditional on other attributes, the coefficients on MPG and size are driven down. We can attempt to investigate the accuracy of this story by including ln(sales) in the cost function keeping in mind that for foreign cars this is not necessarily well measured. (Note, though, in Table 2 that about 80% of the cars in our sample are domestic.) When we include ln(q), so that the cost function is given by,

$$\ln(c_j) = w_j \gamma_w + \gamma_q ln(q_j) + \omega_j,$$

and re-estimate with the same instruments, all cost shifters are positive and significantly different from zero. These estimates are presented in the last two columns of Table 6. The coefficient on ln(q) is very significantly negative although the implied returns to scale seem implausibly high. Adding higher order terms in ln(q) reduces this problem, but we were hesitant to take this approach too far since the data are inaccurate for about a fifth of our sample.

Table 6 does not report the estimated variance covariance matrix of the joint distribution of (ξ, ω), the demand and cost side unobservables. Our estimate of the variance of ω implies that it accounts for about 22% of the estimated variance in log marginal cost. Thus though our estimates do imply that there are some differences in "productivity" across firms, most of the differences in (the log of) marginal costs can be accounted for by a simple linear function of observed characteristics. As one might expect, the correlation between ξ and ω is positive implying that products with more unmeasured quality were more costly to produce. On the other hand, that correlation was only .17, implying that most of the (substantial) variance in ξ could not be accounted for by a linear function of differences in *marginal* costs of production.

Before discussing the demand-side coefficients in the first 3 panels of table 6, we briefly review the structure of purchases in a discrete-choice model. Recall that these are driven by the maximum, and not by the mean, of the utilities associated with the given products. Thus there are, in general, two ways to explain why, say, products with high levels of horsepower to weight (HPWT), are popular. One can explain this by either positing a high mean for the distribution of tastes for HPWT, or by positing a large variance of that same distribution, for both an increase in the mean and an increase in the variance of tastes will increase the share of consumers who purchase cars with high HPWT. However, the two explanations (high average tastes versus a high dispersion of tastes) have different implications for substitution patterns, and thus different implications for how market share will change as product attributes and prices change. If there were, for example, a zero standard deviation for the distribution of marginal utilities of HPWT, we would find that when a high HPWT car increases its price consumers which substitute away from that car have the same marginal utilities for HPWT as any other consumer and hence will not tend to substitute disproportionately toward other high HPWT cars. If, on the other hand, the standard deviation of HPWT was relatively large, the consumers who substitute away from the high HPWT cars will tend to be consumers who placed a high marginal utility on HPWT originally, and hence should tend to substitute disproportionately towards other high HPWT cars [5] .

We now move on to the estimates of the means, $\bar{\beta}_k$, and the standard deviations, σ_k, of the marginal utility distributions. For expositional simplicity, we will focus on the estimates in the first two columns. The demand-side estimates in the decreasing returns to scale case

[5] This same reasoning leads to an interesting set of questions regarding the nonparametric identification of the parameters of the taste distribution (a set of questions which we have not yet begun to investigate). We should note, however, that we had much more difficulty estimating separate mean and variance terms from a single cross section than we did from the panel; indeed the fact that we did not seem to be able to separate out means from variances of the taste distributions from a single cross section was what initially motivated us to go to a panel data set.

imply elasticities and substitution patterns similar to the constant returns to scale case. We find that the means ($\tilde{\beta}$'s) on HP/Weight, Air, and Size are positive and are estimated precisely enough to be significant at traditional significance levels. The estimate of the constant is precise and negative while the mean utility associated with MP$ is negative and insignificantly different from zero. On the other hand, the estimate of the standard deviations of the distribution of marginal utilities of miles per dollar is substantial and it is estimated precisely enough to be considered significant at any reasonable significance level. Thus each of the included attributes is estimated to have either a significantly positive effect on the mean of the distribution of utilities, or a significant positive effect on the standard deviation of that distribution (and in the case of HP/weight, size, and air on both). We turn next to providing some figures on the economic magnitude of these effects.

Table 7 presents estimates of elasticities of demand with respect to attributes and prices. Each row in this table corresponds to a model. The top number in each cell is the actual value of that attribute for that model, while the bottom number is the elasticity of demand with respect to the attribute. For example, the Mazda 323 has a HP/weight ratio of .366 and its elasticity of market share with respect to HP/weight is 0.458.

There are no clear patterns when examining the elasticity of demand with respect to HP/weight. HP/weight is a good proxy for acceleration and we interpret it as such here. The largest cars in the sample, the Lincoln, Cadillac, Lexus, and BMW, have very small elasticities of demand with respect to HP/weight. For each, a 10 percent increase in HP/weight results in less than a 1 percent increase in demand. On the other hand, it appears that consumers who purchase the smallest cars in the sample value acceleration more. For the Mazda 323, Sentra, and Escort, a 10 percent increase in HP/weight increases demand by about 4.5 percent. The relationship between the elasticities and the value of HP/weight is not monotonic though. For mid-size cars, the elasticities are varied. The Maxima (a fairly sporty mid-size car) has a relatively high elasticity (.322) while the similarly sized but more sedate Taurus has an elasticity of (.180). This is consistent with the notion that consumers of sportier cars value HP/weight more.

The effect of having air conditioning standard is estimated to be positive for all cars with this feature and zero for those without. We interpret as implying that consumers who purchase cars with standard air conditioning have a greater preference for luxury.

The elasticities with respect to MP$ illustrate the importance of considering both the mean and standard deviation of the distribution of tastes for a characteristic. The results here are quite intuitive. The elasticity of demand with respect to MP$ declines almost monotonically with the cars MP$ rating. While a 10 percent increase in MP$ increases sales of the Mazda 323, Sentra, and Escort by about a whopping 10 percent, the demand for the cars with low MP$ are actually falling with an increase in MP$. The decreases, though, are quite close to zero. Hence, we conclude that consumers who purchase the high mileage cars care a great deal about fuel economy while those who purchase cars like the BMW 735i or Lexus LS400 basically are not concerned with fuel economy.

The demand elasticities with respect to size are generally declining as cars get larger. This is consistent with the idea that consumers who buy small economy cars would derive substantial utility from a larger car, while those purchasing much larger cars would not derive much extra utility from yet larger autos.

The term on $\ln(y - p)$, α, is of the expected sign and is measured precisely enough to be highly significant. Its magnitude is most easily interpreted by examining the elasticities and markups it, together with the other estimated coefficients, imply. We note first that the estimates imply that demands for all 2217 models in our sample are elastic. The last column of Table 7 lists prices and price elasticities of demand for our sub-sample of 1990 models. We find that the most elastically demanded products are those that are in the most "crowded" market segments – the compact and subcompact models. (The Buick Century is an exception to this pattern.) The Sentra and Mazda 323 face demand elasticities of 6.4 and 6.5 respectively, while the $37,490 BMW and $27,544 (in 1983 dollars) Lexus face demand elasticities of 3.5 and 3.0 respectively.

Table 8 follows Table 5 in presenting a sample of own and cross price semi-elasticities. Each semi-elasticity gives the percentage change in market share for a $1000 increase in price. Looking down the first column, for example, we note that a thousand dollar increase in the price of a Mazda 323 increases the market share of a Nissan Sentra by .705 percent but has almost no effect on the market share of a Lincoln Town Car, Cadillac Seville, Lexus LS400, or a BMW 735i. This table is strikingly different from Table 4 in that it shows cross price elasticities that are large for cars with similar characteristics. Perhaps not surprisingly, the magnitudes of the effects of a \$1000 price increase of the higher priced cars are much smaller than they are for the lower priced cars. The general pattern of cross-price semi-elasticities accords well with intuition. For example, the Lexus is the closest substitute (measured by magnitude of cross price semi-elasticities) to the BMW 735, the Cadillac is the closest substitute to the Lincoln, and the Accord is the closest substitute to the Taurus. Since the demand elasticities will play a crucial role in policy analysis, the sensible elasticities in Table 8 are encouraging.

We are interested in the substitutability of our auto models with the "outside good." Thus, we also calculated ds_0/dp_j. To give some idea of the magnitude of this derivative, we express it as a percentage of the absolute value of the own-price derivative; that is, we calculate

$$\frac{100 * (ds_0/dp_j)}{|ds_j/dp_j|}$$

For a small increase in the price of product j, this gives the number of consumers who substitute from j to the outside good, as a percentage of the total number of consumers who substitute away from j. The results of this exercise are given in table 9. There we report results concerning substitution to the outside good for our sub-sample of 1990 models under both the logit and the BLP specifications.

The first column in table 9 indicates that for every model, about 90 percent of the consumers who substitute away from a model opt instead for the outside good. This figure is just $s_0/(1-s_j)$. The results under the BLP specification are not nearly as uniform across models. Here, the numbers still seem a bit large to us. However, they are much smaller than the corresponding figures for the logit model. The results also show the expected pattern that consumers of lower priced cars are more likely to stay with the outside good when the price of their most preferred model increases. We return to a discussion of modelling the outside alternative in our conclusions.

Table 10 presents the estimated price-marginal cost mark-ups implied by the estimates of the constant returns to scale case reported in Table 6. In 1990, the average markup is \$3,753 and the average ratio of markup to retail price is 0.239.[6] The pattern and magnitudes of the mark-ups reported in table 10 are quite plausible. The models with the lowest markups are the Mazda (\$801), Sentra (\$880), and Escort (\$1077). At the other extreme, the Lexus and BMW have estimated mark-ups of \$9,030 and \$10,975 respectively. In general, mark-ups rise almost monotonically with price.

In the third column of table 10, we list variable profits for each model (since marginal costs are assumed to be constant in output, variable profits are just sales multiplied by price minus marginal cost). Given our estimates, large mark-ups do not necessarily mean large profits, as the sales of some the high markup cars are quite small. The models that, according to our estimates, are the most profitable (by a factor of two, relative to the other models reported in the table) are the Honda Accord and the Ford Taurus. Both are widely regarded as essential to each firm's financial well-being.

A message that falls from Tables 4 through 9 is that allowing more flexible utility specifications gives us a more realistic picture of equilibrium in the *U.S* automobile industry. Conditional on allowing for a more flexible utility specification, there are, however, a number of different variables one might include in the utility and cost functions. It is reasonable to ask how sensitive our results are to our admittedly *ad hoc* choice of included variables. Table 11 begins to address this issue.

There are many ways one might summarize the results of the estimated utility and cost functions. We choose to report the estimated price-marginal cost mark-ups that result from alternative specifications, since these mark-ups embody information from both the cost and demand sides of the model, and they are easily interpretable. The first column of Table 11 replicates the results in Table 10 and is included to make comparisons more convenient. In the second column, we report the mark-ups that result when we include the natural log of output in the cost function. The vector of other cost-shifters, w, is unchanged from the base case. This is the specification reported in the last 2 columns of Table 6 and, as previously noted, the quantity variable is problematic. Nonetheless, the markups follow the same pattern in the base case. The main difference is that the markups

6 Interestingly, while the pattern of mark-ups differs considerably between the logit case and the BLP specification, the average level of markups is similar across the two sets of results.

are uniformly higher. This results from the decreasing returns to scale. The markups over average variable cost (not reported) are much lower. Without higher order terms in ln(q) entering the cost function, the markups over average variable cost are implausibly low. Of all the alternate specifications we investigated, this one yielded the highest price-marginal cost markups, and yet even these markups are not extraordinarily high. For this and all the other alternate specifications, we also report the number of demand side variables whose means or σ's are significantly different from zero at standard levels.

In the results reported in Table 6, the σ associated with air conditioning was not significantly different from zero. We believe the AIR variable is proxying for a degree of luxury. It is possible that there really is little disagreement in the population about this attribute. On the other hand, perhaps it is a poor proxy. In column 3 of Table 11, we report the mark-ups that result from using another proxy – whether automatic transmission is standard equipment. The pattern and magnitudes of the markups are quite similar to the base case results. Markups are slightly higher for the less expensive cars and slightly lower for the high-end cars. It appears our results are robust to a plausible substitute for the proxy for luxury.

In the fourth column of table 11, we report the results when instead of entering the ratio of horsepower to weight, each enter separately and linearly. Of all the alternative specifications investigated, this one gave the largest change in estimated markups. While the patterns of markups is the same, this specification gave implausibly low markups for the more costly cars. This might result if cost were not linear in horsepower and weight, since these cars have large values of each attribute, hence forcing marginal cost to be higher than it perhaps actually is.

In our model, adding additional terms to the cost function is computationally cheap, while adding additional demand side random coefficients is computationally demanding. In column 5 of table 11, we include interaction terms in the cost function between all the continuous characteristics. This captures the notion that the cost of a characteristic may depend on the level of another characteristic. The results of this exercise give markups very similar to our base case results. For most models, the markups are within a few percent of one another. We found that most interaction terms were statistically significant at the usual levels and the elasticities of marginal cost with respect to the continuous attributes were virtually identical to those that resulted with no interaction terms in the cost function. Further, one of the five demand side variables was no longer significant.

In the sixth column of Table 11, we report the mark-ups that result when we replace the constant in the utility function with a set of dummy variables indicating whether the car was built by a firm from the U.S., Japan, or Europe. We also include a variable proxying reliability. The proxy is the reliability rating given by Consumer Reports. This variable is problematic for the reasons given in the discussion of the data. We include it in this particular specification because we suspect that the region dummies may be highly correlated with reliability. If we did not include a measure of reliability, it would mean that an instrument would be correlated with the unobservables, contrary to the assumptions we need for the consistency of our estimator. In this specification there is a separate mean and variance for the dummy associated with each region. Once again, the markups exhibit the same pattern as in the other specifications. We do find, though, that the markups for a number of the models in the middle of the price range are substantially higher. Since these models are not from just one region, it is not clear what is driving this change.

In the final column of Table 11, we report the results when we add weight to the list of regressors (instead of sufficing with the ratios of horsepower to weight), and then allow for interactions in all the cost side variables. Here the linear coefficient of the weight variable came in insignificant on the cost side with a significant mean and insignificant standard deviation on the demand side. The pattern and magnitude of markups was quite similar to the base case results.

7. Applications, Problems, and Extensions.

7.1 Applications.

Our model is defined in terms of four primitives and a Nash equilibrium assumption in prices. The primitives are the utility surface which assigns values to different possible combinations of product characteristics as a function of consumer characteristics, a cost function which determines the production cost associated with different combinations of product characteristics, a distribution of consumer characteristics, and a distribution of product characteristics. Conditional on these primitives (and perhaps a selection criteria to choose among equilibria) the model can solve for the distribution of prices, quantities, variable profits, and consumer welfare. There are, therefore, at least two ways one might use the estimates obtained from it. One is to investigate the impacts of changes in one of the primitives *assuming* that the others are held fixed, while the other is to determine the extent that changes in the various primitives can account for historical movements in the data. The first corresponds to traditional policy analysis, while the second provides an interpretation to the changes that have occurred in the industry over the sample period. In work in progress, we are considering both of these.

It is not difficult to list policy questions that estimates from our model might be used to help analyze. Focusing on some that seem particularly relevant for the auto industry, these include; trade policy issues (the effect of import restrictions), environmental policy issues (carbon and gas guzzler taxes as well as Auto Emission and Corporate Average Fuel Efficiency Standards), merger policy, and the profitability of introducing new cars with alternative vectors of characteristics. Demand elasticities play a crucial role in the analysis of any of these issues. To the extent that estimates of these are not believable, the estimated equilibrium will be suspect. The results in section 6 suggest that the methods developed in this paper give much more realistic estimates of demand elasticities than the more traditional models used to date. On the other hand, all of the models, including our own, are limited by the fact that they provide only a "conditional" analysis of each issue. That is, to use them to do policy analysis we will have to perturb a small number of parameters of the problem and compute new equilibria conditional on an assumption that the rest of the primitives remain unchanged. In many contexts one might not be able to get a realistic account of the effects of the policy change without endogenizing the response of the other primitives to it.

To illustrate, consider the effect of a gas tax. This will induce an immediate change in the demand system (via the role of fuel efficiency in the utility function), and this change in the structure of demand will, in turn, feed through to changes in the relative prices of the models marketed rather quickly. Given appropriate parameter estimates, these are two effects whose impacts on quantities, profits, prices and welfare our model *is* able to analyze (again, perhaps conditional on a selection rule to choose among multiple equilibria).

However, a change in the price of gas will also induce producers to both, introduce more fuel efficient models, and direct research and development activity to decreasing the marginal cost associated with producing models with a higher MPG rating. Thus, in a longer run, both the characteristics of the models marketed, and the characteristics of the cost surface, will evolve as a result of the gas tax, and any analysis which ignores this is likely to be misleading. Similarly, the change in the characteristics of the new cars sold induced by the gas tax will, again in the longer run, induce a change in the characteristics of the used car stock, and this, in turn will feed into the distribution of values of the "outside" alternative, changing, thereby, the distribution of consumer characteristics in the population. In short, what our model misses is both producer and consumer dynamics, and this makes using it to do any longer run policy analysis hazardous. This explains our focus on dynamics in our discussion of extensions below. On the other hand, one has to start somewhere, and any relevant dynamic analysis of producer behavior is likely to require a current profit function, such as the one estimated here. (See for e.g., Maskin and Tirole,1988a and 1988b; and the computational techniques developed to enable us to apply their concept of Markov Perfect Nash Equilibrium to more complicated settings in Pakes and McGuire,1991.

7.2 *Extensions.*

Our methods have been developed around the premise that consumer and producer level data is not always available. This seems an important concession to the realities facing empirical researchers investigating many, but not all, markets. We do note that information on the distribution of many of the relevant consumer characteristics is generally available and illustrate how to make use of the empirical distribution of this information in the estimation algorithm. (In addition to income, consumer characteristics which might be expected to interact with product attributes and for which distributional information is available include household size, geographic region in which the household resides, and age of head of household).

There are however several industries in which some consumer and/or producer-level micro data is available, and the auto industry is one of them. Though production costs for autos are not publicly available at the product-level, the Longitudinal Research Data (LRD) maintained by the Bureau of the Census does contain plant-level cost data. Since industry publications link automotive models to specific plants, we are exploring the possibility of using this information to improve our estimates. Note that separate information on costs would allow for a more detailed examination of the relationship of prices to marginal costs, and, therefore, for a more detailed analysis of the nature of the appropriate equilibrium in the spot market for current output. The cost information would also enable a more flexible analysis of functional forms for the cost surface, and, perhaps, an analysis of how that surface has changed over time in response to changes in both (*R&D*) investments, and in policies.

There is also consumer survey information on automobile purchases (*e.g.* the Survey of Consumer Finance and the Consumer Expenditure Survey). Though these surveys generally do not have products available at the same level of detail as our definition of a "model", and models with small market shares may be owned by no consumer in even a moderately sized survey, there is no doubt that we could obtain both a more detailed, and a more reliable, picture of the demand side of our problem by integrating the survey level information with the aggregate data analyzed in this paper.

The other, perhaps more important, and certainly more difficult, direction for future work is incorporating a realistic treatment of dynamics. On the producer side there are really two aspects of this problem. The first, and definitely easier, is generating an algorithm that provides consistent estimates of the parameters of the static side of the problem that allows for the fact that the unobserved characteristics of an auto are, in part, determined by the same decision process that generates the observed characteristics of the auto, and hence are unlikely to be mean independent of them. The second, and richer, part of the problem is to endogenize the actual choice of the characteristics of the models marketed. As noted earlier, the natural concept of equilibrium in this industry is probably Markov Perfect Nash in investment strategies (see Maskin and Tirole, 1988a and *b*), but even the more detailed Markov Perfect Nash models available to date (see, for example, Ericson and Pakes,1989), still have to be enriched before we can provide a realistic approximation to the multiproduct, multicharacteristic, nature of the auto industry.

Though integrating a complete model of dynamic decision making by consumers (one that incorporates both the transaction costs of buying and selling a car and uncertainty) with a model of producer behavior that is rich enough to encompass the development of new car models is beyond the scope of our current research plans, there are undoubtedly steps in this direction that would provide us with a much more adequate description of consumer behavior while still maintaining manageability. Probably the most important of these in the context of the new car market is providing a more detailed treatment of the consumers' outside alternatives. For many consumers, this is simply driving a used car. Thus, to treat the outside alternative in a realistic way, we will have to bring in the used car market, and we are currently in the process of extending our model in this direction. This is clearly going to be essential to the policy analysis of questions that are primarily related to the characteristics of the stock of cars in use (as most of these are used cars), and to the analysis of policies that differentiate between new and used cars (see, for e.g., Bresnahan and Yao (1985).)

Appendix 1: Consistency and Asymptotic Normality.

This appendix provides arguments which insure the consistency and the asymptotic normality of our estimator. Neither argument requires convergence of $G_J(\theta)$ at values of θ different from θ_0. As a result we do not have to be too detailed about the nature of the dependence induced by the pricing equilibrium. The consistency result does require, however, a condition which places a stochastic bound on the difference between the value of the objective function at a θ different from θ_0, and its value at θ_0. This plays the role of an identification condition, and, like standard identification conditions for nonlinear models, it is only easy to verify for simple special cases (we give an example below). The asymptotic normality results relies on continuity and the asymptotic normality of the objective function at $\theta = \theta_0$ (the latter follows directly from the conditional moment restrictions which underlie the estimation procedure).

Consistency.

Assumption 1.

$\forall \delta > 0, \exists C(\delta)$ such that

$$\lim_{J\to\infty} \Pr \{\inf_{\|\theta-\theta_0\|\geq\delta} \|G_J(\theta) - G_J(\theta_0)\| \geq C(\delta)\} = 1.$$

Proposition 1 .

A1 together with the conditional moment restriction in 4.1 implies that $\theta_J - \theta_0 = o_p(1)$.

Proof.

The conditional moment restriction together with the law of large numbers for triangular arrays (see, for eg. Billingsley, 1979, theorum 6.2) imply that $\|G_J(\theta_0)\| = o_p(1)$. Consequently by theorum 3.1 of Pakes and Pollard (1989) it will suffice to show that for every $(\delta, \epsilon) > (0,0)$ there exists a $C^*(\delta) > 0$ and a $J(\epsilon)$ such that for $J \geq J(\epsilon)$

$$(1)\ \Pr \{\inf_{\|\theta-\theta_0\|\geq\delta} \|G_J(\theta)\| \geq C^*(\delta)\} \geq 1 - \epsilon.$$

From the triangle inequality

$$(2)\ \inf_{\|\theta-\theta_0\|\geq\delta} \|G_J(\theta) - G_J(\theta_0)\| \geq C(\delta)\} => \inf_{\|\theta-\theta_0\|\geq\delta} \|G_J(\theta)\| \geq C(\delta) - \|G_J(\theta_0)\| \ .$$

Fix $\epsilon > 0$, and let $\epsilon^* = \min\{\epsilon, C(\delta)\}$, so that $0 < \epsilon^* \leq \epsilon$. Since $\|G_J(\theta_0)\| = o_p(1)$, $\exists J_1(\epsilon^*)$ such that for any $J \geq J_1(\epsilon^*)$, Pr $\{\|G_J(\theta_0)\| \geq \epsilon^*/2\} \leq \epsilon^*/2$. From A1, $\exists J_2(\epsilon^*)$ such that for $J \geq J_2(\epsilon^*)$, Pr $\{\inf_{\|\theta-\theta_0\|\geq\delta} \|G_J(\theta) - G_J(\theta_0)\| \geq C(\delta)\} \geq 1 - \epsilon^*/2$. Consequently (2) implies that for $J \geq \max\{J_1(\epsilon^*), J_2(\epsilon^*)\}$

$$\Pr\ \{\inf_{\|\theta-\theta_0\|\geq\delta} \|G_J(\theta)\| \geq C(\delta) - \epsilon^*/2\} \geq 1 - \epsilon* \geq 1 - \epsilon.$$

To complete the proof let $C*(\delta) = C(\delta) - \epsilon^*/2 > 0$. ∎.

To get a feeling for the kind of conditions A1 imposes on the problem we consider the special case of the logit demand system described in section 5.1. Here θ is the k-dimensional vector (β, α), so that if we let $\mathbf{x} = (x, p)$, and z be the vector of instruments used to form moment conditions then

$$G_J(\theta) - G_J(\theta_0) = (J^{-1}\Sigma z_j x_j')(\theta - \theta_0).$$

Consequently, a sufficient condition for A1 will be that for each $\epsilon > 0$ there is an $J(\epsilon)$ such that for any $J \geq J(\epsilon)$ the matrix, $J^{-1}\Sigma z_j x_j'$, will have rank k with probability $\geq 1 - \epsilon$, for then $\inf_{\|\theta-\theta_0\|\geq\delta} \|(J^{-1}\Sigma z_j x_j')(\theta - \theta_0)\| \geq \inf_{\|\theta-\theta_0\|\geq\delta} C\|\theta - \theta_0\| \geq C\delta$, as required. In terms of the pricing problem this requires that the price of a product not be a linear function of that product's demand side attributes. However we know that the solution to the pricing problem in (3.3) generates a pricing function which depends on the characteristics of competitor's, as well as on own characteristics.

Asymptotic Normality.

We make three additional assumptions. Assumption 2 can be directly verified for the models estimated in this paper, while if either Assumption 3 or 4 were not true we would expect to see some indication of that fact in the data.

Assumption 2.

Let $\Theta(\delta) = \{\theta \subset \mathbf{R}^k : \|\theta - \theta_0\| \leq \delta\}$, and $\mathcal{G}(\delta)$ be the set of continuously differentiable functions from $\Theta(\delta)$ to $\mathbf{R}^r$ (in mean square). Then $\exists\delta > 0$ such that $G_J(\theta) \in \mathcal{G}(\delta)$ with probability tending to one. The derivative of $G_J(\theta)$ will be denoted by $D_J(\theta)$.

Assumption 3.

θ_0 is in the interior of Θ.

Assumption 4.

Let $G^J(\theta) = E\ G_J(\theta)$, and $\mathcal{G}^*(\delta)$ be the set of continuously differentiable functions from $\Theta(\delta) \to \mathbf{R}^r$ whose derivative matrix at θ_0, say $D^J(\theta_0)$, has the properties that $\|D^J(\theta_0)\| \leq \kappa_1 < \infty$ and determinant$[D^J(\theta_0)'D^J(\theta_0)] \geq \kappa_2 > 0$. Then $\exists\delta > 0$ such that $G^J(\theta) \in \mathcal{G}^*(\delta)$ with probability tending to one.

If $G^J(\theta)$ converged to $G(\theta)$ uniformly in θ in a region of θ_0, then assumption 4 would be satisfied if $D(\theta)$ were continuous in that region and $D(\theta_0)$ had rank k (the dimension of θ).

Proposition 2.

Provided $(\theta_J - \theta_0) = o_p(1)$, A2, A3, and A4, imply that $\sqrt{J}(\theta_J - \theta_0) \to_d N(0, V)$, with V defined as in (4.5).

Proof.

The conditional moment restriction in 4.1 together with a central limit theorum for triangular arrays (see, for eg., Billingsley, 1979, theorem 27.2) implies that $\sqrt{J}G_J(\theta_0)$ converges to a multivariate normal random vector. Thus Theorem 3.3 in Pakes and Pollard [1989; modified in the obvious way to take account of the fact that $E\ G_J(\theta)$ can depend on J], implies that the fact that $(\theta_J - \theta_0) = o_p(1)$, together with A3 and A4, suffice for proposition 2 provided that for every sequence $\{\delta_J\}$ such that $\delta_J \to 0$ as $J \to \infty$

$$(3)\ \sup_{\|\theta-\theta_0\|\leq\delta_J}\{\|G_J(\theta) - G^J(\theta) - G_J(\theta_0)\|/(\sqrt{J} + \|G_J(\theta)\| + \|G^J(\theta)\|)\} = o_p(1).$$

To see that condition (1) is indeed satisfied note that A2, A3, the mean value theorem, and the triangle inequality imply that for J large enough we have with arbitrarily large probability

$$(4)\ \|G_J(\theta) - G^J(\theta) - G_J(\theta_0)\| \leq \|D_J(\theta_0) - D^J(\theta_0)\|\|\theta - \theta_0\| +$$
$$\|D_J[\theta^*_J(\theta) - D_J(\theta_0)\|\|\theta^*_J(\theta) - \theta_0\| + \|D^J[\theta^{*J}(\theta) - D^J(\theta_0)\|\|\theta^{*J}(\theta) - \theta_0\|\ ,$$

where both $\theta^*_J(\theta)$ and $\theta^{*J}(\theta)$ are in the interval $[\theta, \theta_0]$.

Thus to prove the lemma it will suffice to show that, when divided by $\sqrt{J} + \|G^J(\theta)\|$, the supremum over $\|\theta - \theta_0\| \le \delta_J$, of each of the three terms in (4) is $o_p(1)$. Looking at the first term note that

$$\sup_{\|\theta-\theta_0\|\le\delta_J} \{\|D_J(\theta_0) - D^J(\theta_0)\|\|\theta - \theta_0\|/(\sqrt{J} + \|G^J(\theta)\|)\} \le \|[D_J(\theta_0) - D^J(\theta_0)]\sqrt{J}\|\delta_J.$$

Hence it will suffice to show that for any $\epsilon > 0$, $\exists J(\epsilon)$ such that for $J \ge J(\epsilon)$

$$(5)\ \Pr\{\|[D_J(\theta_0) - D^J(\theta_0)]\sqrt{J}\|\delta_J \le \epsilon\} \ge 1 - \epsilon.$$

Recall that at θ_0 each element of $D_J(\theta_0)$ is just the mean of J *i.i.d.* square integrable random variables. Hence from A3 and the Central Limit Theorem for triangular arrays

$$(6)\ \|[D_J(\theta_0) - D^J(\theta_0)]\sqrt{J}\| = O_p(1).$$

(6) implies that for any $\epsilon \exists [J_1(\epsilon), M(\epsilon)]$ such that for $J \ge J_1(\epsilon)$

$$(7)\ \Pr\{\|[D_J(\theta_0) - D^J(\theta_0)]\sqrt{J}\| \le M(\epsilon)\} \ge 1 - \epsilon.$$

Moreover since $\delta_J \to 0$, $\exists J_2(\epsilon)$ such that for $J \ge J_2(\epsilon)$

$$(8)\ \delta_J \le \epsilon/M(\epsilon).$$

(7) and (8) imply (5) for $J \ge \max[J_1(\epsilon), J_2(\epsilon)]$, as required.

Before moving on to the second term in (4) we note that A4 implies that provided $\|\theta - \theta_0\| \le \delta$ then ,

$$\|G^J(\theta)\| \ge C\|\theta - \theta_0\|,$$

for some $C > 0$ with arbitrarily large probability. Consequently for J large enough

$$\begin{aligned}
&\sup_{\|\theta-\theta_0\|\le\delta_J} \{\|D_J[\theta_J^*(\theta)] - D_J(\theta_0)\|\|\theta_J^*(\theta) - \theta_0\|/(\sqrt{J} + \|G^J(\theta)\|)\} \\
\le &\sup_{\|\theta-\theta_0\|\le\delta_J} \{\|D_J[\theta_J^*(\theta)] - D_J(\theta_0)\|\|\theta_J^*(\theta) - \theta_0\|/(C\|\theta - \theta_0\|)\} \\
\le &\sup_{\|\theta-\theta_0\|\le\delta_J} \|D_J[\theta_J^*(\theta)] - D_J(\theta_0)\|C^{-1},
\end{aligned}$$

where the last inequality follows from the fact that $\|\theta_J^*(\theta) - \theta_0\| \leq \|\theta - \theta_0\|$. The fact that this last term is $o_p(1)$ follows directly from the uniform continuity of $D_J(\cdot)$ and the fact that $\delta_J \to 0$. The analogous steps show that the third term in (4) is also $o_p(1)$, which completes the proof of proposition 2. ∎.

Appendix 2: The Contraction Mapping

In this appendix, we will establish the following theorem and then show that the function defined by (5.8) satisfies the hypotheses of the theorem.

Theorem. Consider the complete metric space $(\mathbf{R}^K, d)$ with $d(x,y) = \|x - y\|$ (where $\|\cdot\|$ is the sup-norm.) Let $f : \mathbf{R}^K \to \mathbf{R}^K$ have the properties

(1) $\forall x \in \mathbf{R}^K$, $f(x)$ is continuously differentiable, with, $\forall j$ and k,

$$\partial f_j(x)/\partial x_k \geq 0$$

and

$$\sum_{j=1}^{K} |\partial f_j(x)/\, \partial x_k| < 1.$$

(2) There is a lower bound to $f_j(x)$, denoted $\underline{x}$.

(3) There is a value, $\bar{x}$, with the property that if for any j, $x_j \geq \bar{x}$, then for some k (not necessarily equal to j), $f_k(x) < x_k$.

Let the set $X = [\underline{x}, \bar{x}]^K$. Define the truncated function, $\hat{f} : X \to X$, as $\hat{f}_j(x) = \min\{f_j(x), \bar{x}\}$. Then, $\hat{f}(x)$ is a contraction of modulus less than one on X.

Proof. We will show that

$\exists \beta < 1$ such that $\forall x$ and $x' \in X$, $|\hat{f}(x) - \hat{f}(x')| \leq \beta |x - x'|$.

To see this, choose any x and x' in X and define the scalar $\lambda = \|x - x'\|$. Consider the j^{th} element of $\hat{f}$, $\hat{f}_j(x)$ and WLOG assume $\hat{f}_j(x') - \hat{f}_j(x) \geq 0$. Then, $x + \lambda \geq x'$ implies

$$\hat{f}_j(x') - \hat{f}_j(x) \leq \hat{f}_j(x+\lambda) - \hat{f}_j(x) \leq f_j(x+\lambda) - f_j(x) = \int_0^{\lambda} [\sum_{j=1}^{K} \partial f_j(x+z)/\partial x_k] dz \leq \beta\lambda,$$

where

$$\beta \equiv \max_j \max_{x \in W} \sum_{j=1}^{K} \partial f_j(x)/\partial x_k$$

and the set W is defined as

$$W = \{y \in \mathbf{R}^K : y = (x+z), x \in X\, z \in \mathbf{R}^1, z \leq (\underline{x} - \bar{x})\}.$$

The second inequality in (A2.1) follows from the fact that $\hat{f}_j(x+\lambda) \leq f_j(x+\lambda)$, while $\hat{f}_j(x) = f_j(x)$. The scalar β exists, as it is the maximum of a continuous function over a compact set. β is the maximum value of the integrand in (A2.1) over the set of $(x+z)$ values that can possibly be reached when $x \in X$ and the scalar z is less than the possible difference between any two points in the set X. The final inequality follows from assumption (1), that the integrand of (A2.1) is less than one.

We have now established that $\hat{f}$ is a contraction of modulus β on X. Therefore, there is a unique fixed point, x_o, to $\hat{f}$ on X and for any x in X, the sequence $\hat{f}^n(x)$ converges to x_o. Assumptions (2)-(3) rule out the existence of fixed points to either f or $\hat{f}$ that are outside the interior of X. Thus, x_o cannot be on the boundary of X, x_o is a fixed point of f and there can be no other fixed point to f. (QED)

We will now show that the function $f(\delta) = \delta + \ln(s) - \ln(s(\delta))$ satisfies the hypotheses of the theorem. The function f is differentiable by the differentiability of the function $s(\delta)$. To check the monotonicity condition of assumption one note that

$$\partial f_j(\delta)/\partial\delta_j = 1 - \frac{1}{s_j}\frac{\partial s_j}{\partial \delta_j} ,$$

while for $k \neq j$,

$$\partial f_j(\delta)/\partial\delta_k = -\frac{1}{s_j}\frac{\partial s_j}{\partial \delta_k} .$$

Both $\partial f_j/\partial\delta_j$ and $\partial f_j/\partial\delta_k$ are positive, as can by shown by differentiating our specific market share function and noting $\frac{\partial s_j}{\partial \delta_k} < 0$ and $0 < \frac{\partial s_j}{\partial \delta_j} < s_j$. Berry (1992) notes that

$\sum_{j=1}^{J} |\partial s_j/\partial\delta_k| < 1,$

which in turn establishes that the derivatives of f sum to less than one, establishing all the conditions of assumption one.

It is easy to find the lower bound for f (assumption 2.) Re-write f as

$$f_j(\delta) = \ln(s_j) - \ln(D_j(\delta)), \text{ where}$$

$$D_j(\delta) \equiv \int \frac{e^{\mu_j}}{1+\sum_i e^{\delta_i+\mu_i}} d\Phi(\mu) .$$

As all of the δ_j approach $-\infty$, $D_j(\delta)$ goes to $\int \sum_i e^{\mu_j} d\Phi(\mu)$. Thus a lower bound for f_j is

$\underline{\delta}_j \equiv \ln(s_j) - \ln(\int \sum_i e^{\mu_j} d\Phi(\mu))$.

This is the value of δ_j that would explain a market share for good j of s_j if all the other market shares (other than the outside good) were equal to zero.

Unfortunately, $f(\delta)$ is increasing in δ_j without bound. Berry (1992) does, however, establish the existence of a value, $\bar{\delta}$, such that if any element of δ is greater than $\bar{\delta}$, then there is some k such that $s_k(\delta) > s_k$. For this k, $f_k(\delta) < \delta_k$, satisfying assumption 3. Berry shows that, for product j, an appropriate upper value is the value of δ_j that would explain the market share of the outside good, s_o, when $\delta_o = 0$ and all the other $\delta_k = -\infty$.

References

Anderson, S., A. De Palma, and F. Thisse (1989) "Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach," Review of Economics Studies, 56, 21-35.

Berkovic, J. and J. Rust (1985) "A Nested Logit Model of Automobile Holdings for One Vehicle Households," Transportation Research, 19B, 275-285.

Berry, S. (1991) "Airport Presence as Product Differentiation," American Economic Review, 80, 394-399.

____ (1992) "Discrete Choice Models of Oligopoly Product Differentiation," Yale University working paper.

Bresnahan, T. (1981) "Departures from Marginal-Cost Pricing in the American Automobile Industry," Journal of Econometrics, 17, 201-227.

______ (1987) "Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War," Journal of Econometrics, 35, 457-482.

Bresnahan, T. and D. Yao (1985) "The Nonpecuniary Costs of Automobile Emission Standards," Rand Journal of Economics, 85, 437-455.

Billingsly, P. (1979), *Probability and Measure*, New York, John Wiley and Sons.

Caplin, A. and B. Nalebuff (1991) "Aggregation and Imperfect Competition: On the Existence of Equilibrium," Econometrica, 59, 26-61.

Chamberlin, G. (1986) "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions," Journal of Econometrics, 305-334.

Crain Communications (various issues) Automotive News Market Data Book. Detroit: Crain.

Das, M. S. Olley, and A. Pakes (1992) "The Structure of Equilibrium in the Consumer Electronics Market," in process.

Ericson, R. and Pakes, A. (1989); "An Alternative Theory of Firm and Industry Dynamics", Discussion Paper #445, Columbia University.

Feenstra, R. and J. Levinsohn (1991) "The Characteristics Approach and Oligopoly Pricing," mimeo.

Griliches, Z. (1971) "Hedonic Price Indices for Automobiles: An Econometric Analysis of Quality change," in Price Indices and Quality Change, ed. Z. Griliches. Cambridge: Harvard University Press.

Hansen, L. (1982) 'Large Sample Properties of Generalized Method of Moments Estimators," Econometrica, 50, 1029-1054.

Haseman, J. and L. Kuper (1979) "Analysis of Dichotomous Response Data from Certain Toxological Experiments," Biometrics, 35, 281-293.

Hausman, J. (1987) "Project Independence Report: An Appraisal of U.S. Energy Needs Up to 1985," Rand Journal of Economics, 517-551.

Koujianou, P. (1991) "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," mimeo.

Lerman, S. and C. Manski (1981) "On the Use of Simulated Frequencies to Approximate Choice Probabilities," in Structural Analysis of Discrete Data with Econometric Applications. Cambridge: MIT Press.

Levinsohn, J. (1988) "Empirics of Taxes on Differentiated Products: The Case of Tariffs in the U.S. Automobile Industry," in Trade Policy Issues and Empirical Analysis, ed. R. Baldwin Chicago: University of Chicago Press.

Manski, C. (1983) "Analysis of Equilibrium Automobile Holdings in Israel with Aggregate Discrete Choice Models," Transportation Research, **17B**, 373-389.

Manski, C. and D. McFadden, Eds. (1981) Structural Analysis of Discrete Data with Econometric Applications. Cambridge: MIT Press.

Maskin, E., and J. Tirole (1988a) "A Theory of Dynamic Oligopoly,I: Overview and Quantity Competition with Large Fixed Costs," Econometrica, Vol 56. pp.549-569.

Maskin, E., and J. Tirole (1988b) "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles" Econometrica, Vol. 56, pp. 571-99.

McFadden, D. (1973) "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers of Econometrics, ed. Zarembka, P. New York: Academic.

______ (1981) "Econometric Models of Probabilistic Choice," in Structural Analysis of Discrete Data with Economctric Applications, ed. C. Manski and D. McFadden, Cambridge: MIT.

______ (1986) "Econometric Analysis of Qualitative Response Models," in Handbook of Econometrics, Volume III, ed. Griliches and Intriligator. Amsterdam: North Holland.

Morgan, M. (1990) *The History of Econometric Ideas*, New York: Cambridge.

Morrison, S. and C. Winston (1989) "Enhancing the Performance of the Deregulated Air Transportation System," Brookings Papers on Micreconomics, 1, 61-112.

Newey, W. (1990) "Efficient Instrumental Variables Estimation of Nonlinear Models," Econometrica, 58, 809-839.

Ohta, M. and Z. Griliches (1976) "Automobile Prices Revisited: Extensions of the Hedonic Hypothesis," in *Household Production and Consumption*, ed. N. Terleckyj, New York: NBER.

Pakes, A. (1986) "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," Econometrica, 54, 755-784.

____ (forthcoming) "Dynamic Structural Models, Problems and Prospects; Mixed Continuous Discrete Controls and Market Interactions," in Advances in Econometrics, ed. J.J. Laffont and C. Sims. : forthcoming.

Pakes, A., and P. McGuire (1991) "Computation of Markov Perfect Nash Equilibria I: Numerical Implications of a Dynamic Differentiated Product Model" N.B.E.R. Technical Working Paper.

Pakes, A., and D. Pollard (1989) "Simulation and the Asymptotics of Optimization

Estimators," Econometrica, pp. 1027-10057.

Perloff, J. and S. Salop (1985) "Equilibrium with Product Differentiation," Review of Economic Studies, **52**, 107-120.

Rosen, S. (1974) "Hedonic Prices and Implicit Markets: Product differentiation in pure competition," Journal of Political Economy, **82**, 34-55.

Rubinstein, R. (1981) *Simulation and the Monte Carlo Method*, New York: Wiley.

Sattinger, M. (1984) "Value of an Additional Firm in Monopolistic Competition," Review of Economic Studies, **49**, 3-13.

Schamlensee, R. (1985) "Econometric Diagnosis of Competitive Localization", International Journal of Industrial Organization, pp.57-70.

Shaked, A. and J. Sutton (1982) "Relaxing Price Competition Through Product Differentiation," Review of Economic Studies, **49**, 3-13.

Suits, D. (1958) "The Demand for Automobiles in the United States," Review of Economics and Statistics, **40**, 273-280.

Toder, E. with N. Cardell and E. Burton (1978) Trade Policy and the *U.S.* Automobile Industry. New York: Praeger.

Trajtenberg, M. (1989) "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," Journal of Political Economy, **97**, 444-479.

Williams, D. (1982) "Extra-binomial Variation in Logistic Linear Models," Applied Statistics, **31**, 144-148.

Working, E. (1926) "What Do Statistical Demand Curves Show?," Quarterly Journal of Economics, **41**, 212-235.

TABLE 1

Standard Deviations of the Simulation Error.

	Frequency	Smooth	Importance
Mean Std. Dev. of shares (one draw)	0.0228	0.0032	0.0009
Mean Std. Dev. of δ (200 draws)	—	1.6711	0.3040

TABLE 2
Descriptive Statistics

Year	No. of Models	Quantity	Price	Domestic	Japan	European	HP/Wt	Size	Air	MPG	MP$
1971	92	86.892	7.868	0.866	0.057	0.077	0.490	1.496	0.000	1.662	1.850
1972	89	91.763	7.979	0.892	0.042	0.066	0.391	1.510	0.014	1.619	1.875
1973	86	92.785	7.535	0.932	0.040	0.028	0.364	1.529	0.022	1.589	1.819
1974	72	105.119	7.506	0.887	0.050	0.064	0.347	1.510	0.026	1.568	1.453
1975	93	84.775	7.821	0.853	0.083	0.064	0.337	1.479	0.054	1.584	1.503
1976	99	93.382	7.787	0.876	0.081	0.043	0.338	1.508	0.059	1.759	1.696
1977	95	97.727	7.651	0.837	0.112	0.051	0.340	1.467	0.032	1.947	1.835
1978	95	99.444	7.645	0.855	0.107	0.039	0.346	1.405	0.034	1.982	1.929
1979	102	82.742	7.599	0.803	0.158	0.038	0.348	1.343	0.047	2.061	1.657
1980	103	71.567	7.718	0.773	0.191	0.036	0.350	1.296	0.078	2.215	1.466
1981	116	62.030	8.349	0.741	0.213	0.046	0.349	1.286	0.094	2.363	1.559
1982	110	61.893	8.831	0.714	0.235	0.051	0.347	1.277	0.134	2.440	1.817
1983	115	67.878	8.821	0.734	0.215	0.051	0.351	1.276	0.126	2.601	2.087
1984	113	85.933	8.870	0.783	0.179	0.038	0.361	1.293	0.129	2.469	2.117
1985	136	78.143	8.938	0.761	0.191	0.048	0.372	1.265	0.140	2.261	2.024
1986	130	83.756	9.382	0.733	0.216	0.050	0.379	1.249	0.176	2.416	2.856
1987	143	67.667	9.965	0.702	0.245	0.052	0.395	1.246	0.229	2.327	2.789
1988	150	67.078	10.069	0.717	0.237	0.045	0.396	1.251	0.237	2.334	2.919
1989	147	62.914	10.321	0.690	0.261	0.049	0.406	1.259	0.289	2.310	2.806
1990	131	66.377	10.337	0.682	0.276	0.043	0.419	1.270	0.308	2.270	2.852
all	2217	78.804	8.604	0.790	0.161	0.049	0.372	1.357	0.116	2.099	2.086

Notes: The entry in each cell is the sales weighted mean.

TABLE 3
The Range of Continuous Demand Characteristics
(and associated models)

Variable	Percentile 0	25	50	75	100
Price	90 Yugo 3.393	79 Mercury Capri 6.711	87 Buick Skylark 8.728	71 Ford T-Bird 13.074	89 Porsche 911 Cabriolet 68.597
Sales	73 Toyota 1600CR .049	72 Porsche Rdstr 15.479	77 Plym. Arrow 47.345	82 Buick LeSabre 109.002	71 Chevy Impala 577.313
HP/Wt.	85 Plym. GranFury 0.170	85 Suburu DH 0.337	86 Plym. Caravelle 0.375	89 Toyota Camry 0.428	89 Porsche 911 Turbo 0.948
Size	73 Honda Civic 0.756	77 Renault GTL 1.131	89 Hyundai Sonata 1.270	81 Pontiac F-Bird 1.453	73 Im. Imperial 1.888
MP$	74 Cad. Eldorado 8.46	78 Buick Skyhawk 15.57	82 Mazda 626 20.10	84 Pontiac 2000 24.86	89 Geo Metro 64.37
MPG	74 Cad. Eldorado 9	79 BMW 528i 17	81 Dodge Challenger 20	75 Suburu DL 25	89 Geo Metro 53

Notes: The top entry for each cell gives the model name and the number directly below it gives the value of the variable for this model.

TABLE 4
Results with Logit Demand and Marginal Cost Pricing
(2217 observations)

	OLS Logit Demand	IV Logit Demand	OLS $ln(price)$ on w
Variable			
Constant	-10.068	-9.273	1.882
	(0.253)	(0.493)	(0.119)
HP/Weight *	-0.121	1.965	0.520
	(0.277)	(0.909)	(0.035)
Air	-0.035	1.289	0.680
	(0.073)	(0.248)	(0.019)
MP$	0.263	0.052	-
	(0.043)	(0.086)	
MPG*	-	-	-0.471
			(0.049)
Size *	2.341	2.355	0.125
	(0.125)	(0.247)	(0.063)
trend	-	-	0.013
			(0.002)
Price	-0.089	-0.216	-
	(0.004)	(0.123)	
No. Inelastic Demands (+/- 2 s.e.'s)	1494 (1429-1617)	22 (7-101)	n.a.
R^2	0.387	n.a.	.656

Notes: The standard errors are reported in parentheses.

* The continuous product characteristics – hp/weight, size and fuel efficiency (MP$ or MPG) – enter the demand equations in levels, but enter the column 3 price regression in natural logs.

TABLE 5

A Sample from 1990 of Estimated Own- and Cross-Price Semi-Elasticities:
Based on Table 4 IV Estimates

	Mazda 323	Nissan Sentra	Ford Escort	Chevy Cavalier	Honda Accord	Ford Taurus	Buick Century	Nissan Maxima	Acura Legend	Lincoln Town Car	Cadillac Seville	Lexus LS400	BMW 735i
323	21.596	0.0112	0.0653	0.0668	0.0944	0.0709	0.0254	0.0226	0.0121	0.0336	0.0075	0.0093	0.0024
Sentra	0.0052	21.589	0.0653	0.0668	0.0944	0.0709	0.0254	0.0226	0.0121	0.0336	0.0075	0.0093	0.0024
Escort	0.0052	0.0112	21.535	0.0668	0.0944	0.0709	0.0254	0.0226	0.0121	0.0336	0.0075	0.0093	0.0024
Cavalier	0.0052	0.0112	0.0653	21.533	0.0944	0.0709	0.0254	0.0226	0.0121	0.0336	0.0075	0.0093	0.0024
Accord	0.0052	0.0112	0.0653	0.0668	21.506	0.0709	0.0254	0.0226	0.0121	0.0336	0.0075	0.0093	0.0024
Taurus	0.0052	0.0112	0.0653	0.0668	0.0944	21.529	0.0254	0.0226	0.0121	0.0336	0.0075	0.0093	0.0024
Century	0.0052	0.0112	0.0653	0.0668	0.0944	0.0709	21.575	0.0226	0.0121	0.0336	0.0075	0.0093	0.0024
Maxima	0.0052	0.0112	0.0653	0.0668	0.0944	0.0709	0.0254	0.0226	0.0121	0.0336	0.0075	0.0093	0.0024
Legend	0.0052	0.0112	0.0653	0.0668	0.0944	0.0709	0.0254	21.577	21.588	0.0336	0.0075	0.0093	0.0024
TownCar	0.0052	0.0112	0.0653	0.0668	0.0944	0.0709	0.0254	0.0226	0.0121	21.566	0.0075	0.0093	0.0024
Seville	0.0052	0.0112	0.0653	0.0668	0.0944	0.0709	0.0254	0.0226	0.0121	0.0336	21.592	0.0093	0.0024
LS400	0.0052	0.0112	0.0653	0.0668	0.0944	0.0709	0.0254	0.0226	0.0121	0.0336	0.0075	21.591	0.0024
735i	0.0052	0.0112	0.0653	0.0668	0.0944	0.0709	0.0254	0.0226	0.0121	0.0336	0.0075	0.0093	21.598

Notes: Cell entries i, j, where i indexes row and j column, give the percentage change in market share of i with a \$1000 change in the price of j.

TABLE 6
Estimated Parameters of the Demand and Pricing Equations:
BLP Specification, 2217 observations

	Variable	Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
Demand Side Parameters					
Means ($\bar{\beta}$'s)	Constant	-7.061	0.752	-7.304	0.698
	HP/Weight	2.883	0.972	2.185	0.743
	Air	1.521	0.816	0.579	0.662
	MP$	-0.122	0.240	-0.049	0.208
	Size	3.460	0.510	2.604	0.357
Std. Deviations (σ_β's)	Constant	3.612	1.520	2.009	1.041
	HP/Weight	4.628	1.136	1.586	1.270
	Air	1.818	1.687	1.215	1.290
	MP$	1.050	0.213	0.670	0.203
	Size	2.056	0.542	1.510	0.336
Term on Price (α)	$\ln(y-p)$	43.501	5.254	23.710	3.289
Cost Side Parameters					
	Constant	0.952	0.151	0.726	0.237
	ln(HP/Weight)	0.477	0.039	0.313	0.056
	Air	0.619	0.073	0.290	0.045
	ln(MPG)	-0.415	0.052	0.293	0.090
	ln(Size)	-0.046	0.066	1.499	0.150
	Trend	0.019	0.002	0.026	0.003
	$ln(q)$			-0.387	0.036

TABLE 7
A Sample from 1990 of Estimated Demand Elasticities
With Respect to Attributes and Price
(Based on Table 6 (CRTS) Estimates)

Model	Value of Attribute/Price Elasticity of demand with respect to:				
	HP/Weight	Air	MP$	Size	Price
Mazda323	0.366	0.000	3.645	1.075	5.049
	0.458	0.000	1.010	1.338	6.358
Sentra	0.391	0.000	3.645	1.092	5.661
	0.440	0.000	0.905	1.194	6.528
Escort	0.401	0.000	4.022	1.116	5.663
	0.449	0.000	1.132	1.176	6.031
Cavalier	0.385	0.000	3.142	1.179	5.797
	0.423	0.000	0.524	1.360	6.433
Accord	0.457	0.000	3.016	1.255	9.292
	0.282	0.000	0.126	0.873	4.798
Taurus	0.304	0.000	2.262	1.334	9.671
	0.180	0.000	-0.139	1.304	4.220
Century	0.387	1.000	2.890	1.312	10.138
	0.326	0.701	0.077	1.123	6.755
Maxima	0.518	1.000	2.513	1.300	13.695
	0.322	0.396	-0.136	0.932	4.845
Legend	0.510	1.000	2.388	1.292	18.944
	0.167	0.237	-0.070	0.596	4.134
TownCar	0.373	1.000	2.136	1.720	21.412
	0.089	0.211	-0.122	0.883	4.320
Seville	0.517	1.000	2.011	1.374	24.353
	0.092	0.116	-0.053	0.416	3.973
LS400	0.665	1.000	2.262	1.410	27.544
	0.073	0.037	-0.007	0.149	3.085
BMW 735i	0.542	1.000	1.885	1.403	37.490
	0.061	0.011	-0.016	0.174	3.515

Notes: The value of the attribute or, in the case of the last column, price, is the top number and the number below it is the elasticity of demand with respect to the attribute (or, in the last column, price.)

TABLE 8

A Sample from 1990 of Estimated Own- and Cross-Price Semi-Elasticities:
Based on Table 6 (CRTS) Estimates

	Mazda 323	Nissan Sentra	Ford Escort	Chevy Cavalier	Honda Accord	Ford Taurus	Buick Century	Nissan Maxima	Acura Legend	Lincoln Town Car	Cadillac Seville	Lexus LS400	BMW 735i
323	-125.933	1.518	8.954	9.680	2.185	0.852	0.485	0.056	0.009	0.012	0.002	0.002	0.000
Sentra	0.705	-115.319	8.024	8.435	2.473	0.909	0.516	0.093	0.015	0.019	0.003	0.003	0.000
Escort	0.713	1.375	-106.497	7.570	2.298	0.708	0.445	0.082	0.015	0.015	0.003	0.003	0.000
Cavalier	0.754	1.414	7.406	-110.972	2.291	1.083	0.646	0.087	0.015	0.023	0.004	0.003	0.000
Accord	0.120	0.293	1.590	1.621	-51.637	1.532	0.463	0.310	0.095	0.169	0.034	0.030	0.005
Taurus	0.063	0.144	0.653	1.020	2.041	-43.634	0.335	0.245	0.091	0.291	0.045	0.024	0.006
Century	0.099	0.228	1.146	1.700	1.722	0.937	-66.635	0.773	0.152	0.278	0.039	0.029	0.005
Maxima	0.013	0.046	0.236	0.256	1.293	0.768	0.866	-35.378	0.271	0.579	0.116	0.115	0.020
Legend	0.004	0.014	0.083	0.084	0.736	0.532	0.318	0.506	-21.820	0.775	0.183	0.210	0.043
TownCar	0.002	0.006	0.029	0.046	0.475	0.614	0.210	0.389	0.280	-20.175	0.226	0.168	0.048
Seville	0.001	0.005	0.026	0.035	0.425	0.420	0.131	0.351	0.296	1.011	-16.313	0.263	0.068
LS400	0.001	0.003	0.018	0.019	0.302	0.185	0.079	0.280	0.274	0.606	0.212	-11.199	0.086
735i	0.000	0.002	0.009	0.012	0.203	0.176	0.050	0.190	0.223	0.685	0.215	0.336	-9.376

Notes: Cell entries i, j, where i indexes row and j column, give the percentage change in market share of i with a \$1000 change in the price of j.

TABLE 9
Substitution to the Outside Good

Model	Given a price increase, the percentage who substitute to the outside good (as a percentage of all who substitute away.)	
	Logit	BLP
Mazda 323	90.870	27.123
Nissan Sentra	90.843	26.133
Ford Escort	90.592	27.996
Chevy Cavalier	90.585	26.389
Honda Accord	90.458	21.839
Ford Taurus	90.566	25.214
Buick Century	90.777	25.402
Nissan Maxima	90.790	21.738
Acura Legend	90.838	20.786
Lincoln TownCar	90.739	20.309
Cadillac Seville	90.860	16.734
Lexus LS400	90.851	10.090
BMW 735i	90.883	10.101

TABLE 10

A Sample from 1990 of Estimated Price–Marginal Cost Markups and Variable Profits: Based on Table 6 (CRTS) Estimates

	Price	Markup over MC $(p - MC)$	Variable Profits (in $'000's) $q * (p - MC)$
Mazda 323	$ 5,049	$ 801	$ 18,407
Nissan Sentra	$ 5,661	$ 880	$ 43,554
Ford Escort	$ 5,663	$ 1,077	$ 311,068
Chevy Cavalier	$ 5,797	$ 1,302	$ 384,263
Honda Accord	$ 9,292	$ 1,992	$ 830,842
Ford Taurus	$ 9,671	$ 2,577	$ 807,212
Buick Century	$ 10,138	$ 2,420	$ 271,446
Nissan Maxima	$ 13,695	$ 2,881	$ 288,291
Acura Legend	$ 18,944	$ 4,671	$ 250,695
Lincoln TownCar	$ 21,412	$ 5,596	$ 832,082
Cadillac Seville	$ 24,353	$ 7,500	$ 249,195
Lexus LS400	$ 27,544	$ 9,030	$ 371,123
BMW 735i	$ 37,490	$ 10,975	$ 114,802

TABLE 11
Results from some Alternative Specifications:
Price-Marginal Cost Markups

	Base Case (reported in Table7)	Include ln(q) in cost fnc.	Use AT instead of AIR	Weight and HP instead of HP/Wt.	Include interaction terms in cost fnc.	Use 3 Region Dummies and add Reliability	Add weight and include interactions in cost fnc.
Mazda 323	$ 801	$ 1,616	$ 1,012	$ 1,073	$ 828	$ 1,125	$ 1,389
Nissan Sentra	$ 880	$ 1,769	$ 1,153	$ 1,271	$ 912	$ 1,308	$ 1,487
Ford Escort	$ 1,077	$ 2,043	$ 1,326	$ 1,470	$ 1,111	$ 2,094	$ 1,690
Chevy Cavalier	$ 1,302	$ 2,490	$ 1,729	$ 1,655	$ 1,329	$ 2,593	$ 2,020
Honda Accord	$ 1,992	$ 3,059	$ 2,629	$ 2,703	$ 2,059	$ 3,839	$ 2,327
Ford Taurus	$ 2,577	$ 3,721	$ 2,528	$ 3,344	$ 2,585	$ 4,094	$ 2,898
Buick Century	$ 2,420	$ 4,162	$ 3,161	$ 2,939	$ 2,405	$ 4,030	$ 3,321
Nissan Maxima	$ 2,881	$ 4,674	$ 4,565	$ 2,085	$ 2,911	$ 6,941	$ 3,513
Acura Legend	$ 4,671	$ 7,105	$ 6,563	$ 3,059	$ 4,661	$ 8,305	$ 5,061
Lincoln TownCar	$ 5,596	$ 8,029	$ 6,778	$ 4,765	$ 5,508	$ 7,114	$ 6,518
Cadillac Seville	$ 7,500	$ 10,733	$ 8,635	$ 4,863	$ 7,439	$ 9,182	$ 8,015
Lexus LS400	$ 9,030	$ 10,510	$ 8,411	$ 4,791	$ 8,585	$ 10,925	$ 7,398
BMW 735i	$10,975	$ 13,646	$ 9,122	$ 7,605	$ 10,713	$ 12,153	$ 12,202
No. of demand side variables significant at 95% level[1]	5 of 5	4 of 5	4 of 5	6 of 6	4 of 5	8 of 8	4 of 6

1. A demand side variable is considered significant if *either* its mean or or standard deviation (σ_β) is significant. See text for details.