Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market

Steven Berry

Yale University and National Bureau of Economic Research

James Levinsohn

University of Michigan and National Bureau of Economic Research

Ariel Pakes

Harvard University and National Bureau of Economic Research

In this paper, we consider how rich sources of information on consumer choice can help to identify demand parameters in a widely used class of differentiated products demand models. Most important, we show how to use "second-choice" data on automotive purchases to obtain good estimates of substitution patterns in the automobile industry. We use our estimates to make out-of-sample predictions about important recent changes in industry structure.

We thank numerous seminar participants, two referees, and the editors Lars Hansen and John Cochrane for helpful suggestions. We also thank the National Science Foundation for financial support, through grants 9122672, 9512106, and 9617887. We are particularly grateful to G. Mustafa Mohatarem at General Motors Corp., who made possible our access to the data, and for further help from GM's Robert Bordley. Gautam Gowrisankaran, Dan Ackerberg, Lanier Benkard, Amil Petrin, and Nadia Soboleva provided invaluable assistance. Now that they are all successful academics, we hope their own research assistants come close to matching their standard. The most recent revision of this paper would not have been possible without Nadia Soboleva's expert advice.

[*Journal of Political Economy,* 2004, vol. 112, no. 1, pt. 1]
 0022-3808/2004/11201-0002$10.00

I. Introduction

In this paper, we consider how rich sources of information on consumer choice can help to identify demand parameters in a widely used class of differentiated products demand models. The demand framework is a class of differentiated product demand models whose foundations date back at least to Lancaster (1972) and McFadden (1974). In these models, products are described as bundles of characteristics, and consumers choose the product that maximizes the utility derived from product characteristics.

We follow in a tradition that seeks to uncover basic parameters of demand and supply so that we can obtain a detailed analysis of past events and make realistic predictions about out-of-sample policies and changes in industry structure. To illustrate, we conclude with an analysis of two of our sample changes: the recent decision of General Motors to shut down its historic Oldsmobile division and the introduction of luxury sport utility vehicles (SUVs). Our data indicate tight substitution patterns between similar products, and so our estimates predict that General Motors will hold on to a substantial fraction of its former Oldsmobile customers. Also, we find significant potential demand for "high-end" SUVs in 1993, consistent with the later introduction of such vehicles.

Our estimates make use of a novel data set, provided to us by General Motors, that surveys recent purchasers of automobiles. The most novel aspect of our data is the presence of consumers' "second choices"—the purchase that they would have made if their preferred product were not available. In our example, we find that this kind of data is very helpful in estimating the model parameters that govern the predicted pattern of substitution across products. The second-choice data are similar to other kinds of survey data on product rankings, although they may be of higher quality because our consumers have recently completed a very expensive and somewhat time-consuming purchase.

In earlier work (e.g., Berry, Levinsohn, and Pakes 1995), we emphasized estimation strategies based on changes across markets (or across time) in the choice set facing consumers. In that work, we assume that the distribution of consumers' underlying tastes, conditional on an observed distribution of consumer incomes and demographics, is invariant across markets/time. We then propose to estimate substitution patterns from data on how choices vary as the characteristics and numbers of products, as well as the distribution of observed consumer attributes, change across markets. Thus, in our earlier paper and related papers, the model parameters that govern substitution patterns are estimated from data on (i) how consumers substitute across products when the characteristics, prices, and number of products change and (ii) how the

distribution of consumer attributes changes choices for a given choice set.

Many authors have also made use of data that match consumer attributes to consumer choices. (This includes most of the early discrete choice demand literature and also recent work in industrial organization by Goldberg [1995] and Petrin [2002].) These data, together with changing choice sets, can help to estimate substitution patterns to the degree that these patterns are explained by observed consumer attributes. For example, Petrin finds that consumer attribute data (together with a dramatically changing choice set) are quite useful in explaining substitution patterns (and welfare results) for minivans.[1]

In the present paper, the second-choice data provide an alternative source of identification. These second-choice data have several strong advantages. First, they give us a direct, data-based measure of substitution. As a result, we can ask what classes of models are capable of reproducing this observed pattern of substitution. For example, we find that models without unobserved heterogeneity (but with observed consumer attributes) do a bad job of reproducing observed substitution patterns. Also, and perhaps more important, by requiring the model parameters to match the observed second-choice substitution patterns, we gain a source of identifying power that does not rely on exogenous changes in choice sets.

We do find, however, the not very surprising result that second-choice data on a single-market cross section of products (without any variation in prices for a given vehicle) cannot by themselves identify the absolute level of price elasticities (as opposed to the pattern of substitution across products). Thus, even high-quality second-choice data will not solve all estimation problems in this class of models. In the context of our single cross section of data, we discuss several ways of bringing information from outside sources to fix the level of price elasticities. This allows us to perform our policy experiments.[2]

In the remainder of this paper, we first review the basic empirical differentiated products demand model from the recent industrial organization literature. We then describe our estimation procedure, emphasizing the role it gives to different sources of data. After describing the data and the parameter estimates, we provide results on the policy experiments.

[1] The result for minivans is consistent with our results as well, but we show that other automotive choices are not as closely tied to commonly observed consumer attributes. Also note that variation in consumer attributes sometimes effectively changes the choice set: if one does not live near public transportation, then it is not really an option.

[2] Future work might focus on combining different sources of information, including the kind of cross-market data that we ourselves used in earlier work.

II. The Model

We start from the model in Berry et al. (1995) (BLP), which is a model of household choice that is then explicitly aggregated to obtain product-level demands. It is therefore able to analyze *both* our micro data on household choices and our aggregate data on product-level demands in one consistent framework.

Largely for simplicity, we use a linear version of the utility, u_{ij}, that consumer i obtains from the choice of product j (this follows the traditional discrete choice random coefficients literature, e.g., Domencich and McFadden [1975] or Hausman and Wise [1978]). Let $j = 0, \ldots, J$ index the products competing in the market, where product $j = 0$ is the "outside" good (so that u_{i0} is the utility of the consumer if she does not purchase any of these J goods and instead allocates all income to other purchases). Let k index the observed (by us) product characteristics, including price, and r index the observed household attributes.

Our model is then

$$u_{ij} = \sum_k x_{jk}\tilde{\beta}_{ik} + \xi_j + \epsilon_{ij}, \tag{1}$$

with

$$\tilde{\beta}_{ik} = \bar{\beta}_k + \sum_r z_{ir}\beta^o_{kr} + \beta^u_k \nu_{ik}, \tag{2}$$

where the x_{jk} and ξ_j are, respectively, observed and unobserved *product* characteristics, the $\tilde{\beta}_{ik}$ represent the "taste" of consumer i for product characteristic k, the $\mathbf{z}_i$ and $\boldsymbol{\nu}_i$ are vectors of observed and unobserved *consumer* attributes, and the ϵ_{ij} represent idiosyncratic individual preferences, assumed to be independent of the product attributes and of each other. Note that the model allows consumers to differ in their tastes for different product characteristics. Those differences (the $\tilde{\boldsymbol{\beta}}$) are allowed (via eq. [2]) to depend on *both* consumer attributes observed by the econometrician (through $\boldsymbol{\beta}^o$, where the o superscript is for "observed") and attributes that the econometrician does not observe (through $\boldsymbol{\beta}^u$, where u is for "unobserved").[3] In our example the $\mathbf{z}$ vectors contain consumer attributes listed in our data (e.g., income, family size, and age of household head), and the $\boldsymbol{\nu}$ vectors allow for consumer attributes that are not in our data (e.g., distance to work or a need to

[3] Equations (1) and (2) make several simplifying assumptions, including that there is only one unobserved product characteristic and consumers do not differ in their preferences for it. These simplifications are not necessary for the arguments that follow, though they simplify both the exposition and the subsequent computations; see Heckman and Snyder (1997) for a related model with a higher dimension of unobserved characteristics and Das, Olley, and Pakes (1994) for an attempt to let consumers differ in their preferences for the unobserved characteristic in this model.

transport a little league team). Similarly, the x_k are auto characteristics that we measure (e.g., price, size, and horsepower) and the ξ are unmeasured aspects of car quality.

We want to stress two features of this framework: the interaction terms and the product-specific constant terms. First, as noted in the earlier literature (see McFadden et al. 1977; Hausman and Wise 1978; Berry et al. 1995), the interaction between consumer tastes and product characteristics determines substitution patterns in discrete choice models. As the variance in the random tastes for product characteristics increases, similar products (in the space of $\mathbf{x}$'s) become better substitutes. Models without individual differences in preferences for characteristics generate demand substitution patterns that are known to be a priori unreasonable (depending only on market shares and not on the characteristics of the vehicles). A goal of this paper is to provide accurate measures of substitution patterns, and so we allow for unobserved (as well as observed) determinants of characteristic preferences.

Second, vehicles (and most other consumer products) are differentiated from one another in many dimensions. We shall include characteristics that proxy for the most important sources of differentiation, but even if we had the data, we could not hope to estimate the distribution of preferences over a set of characteristics that is large enough to capture all aspects of product differentiation. The role of the unobserved product characteristic, ξ, is to pick up the total impact of the characteristics not included in our specification. As stressed in Berry (1994) and in Berry et al. (1995), one might expect ξ to be correlated with price: products with higher unmeasured quality might sell at a higher price. This is the differentiated product analogue of the standard "simultaneity" problem in demand analysis, and our previous work indicates that when we do not account for this correlation, we obtain unreasonably small (in absolute value) price elasticities.

The consumer-level choice model is found by substituting equation (2) into (1) to obtain

$$u_{ij} = \delta_j + \sum_{kr} x_{jk} z_{ir} \beta^o_{kr} + \sum_k x_{jk} \nu_{ik} \beta^u_k + \epsilon_{ij}, \tag{3}$$

where, for $j = 0, 1, \ldots, J$,

$$\delta_j = \sum_k x_{jk} \bar{\beta}_k + \xi_j. \tag{4}$$

This equation clarifies two important points about the identification of our model. First, even without an assumption on the joint distribution of $(\xi, \mathbf{x})$, the micro data allow us to estimate some but *not all* of the parameters of the model. Second, the remaining parameters determine

the elasticities of interest, and identifying these parameters requires assumptions of the sort used in market-level data.

To see that some parameters are identified without assumptions on $(\xi, \mathbf{x})$, note that equation (3) defines a traditional random coefficients discrete choice model with choice-specific constant terms, δ_j. Given parametric assumptions on $(\boldsymbol{\nu}, \boldsymbol{\epsilon})$ and standard regularity conditions, we can therefore obtain consistent estimators of the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\delta}, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u)$ from micro data (such as our CAMIP data) without assumptions about the unobservable ξ's.[4] Some questions of interest require only these parameters. One important example is the calculation of ideal price indices (see Pakes, Berry, and Levinsohn 1993) (Sec. VII contains another example).

However, knowledge of $\boldsymbol{\theta} = (\boldsymbol{\delta}, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u)$ does not identify own and cross-price (and characteristic) elasticities. Unless product characteristics have no systematic effect on demand ($\bar{\boldsymbol{\beta}} \equiv 0$), the choice-specific constant $\boldsymbol{\delta}$ is itself a function of product characteristics. Thus to calculate the impact of, say, price on demand, we need to know the impact of price on $\boldsymbol{\delta}$; that is, we need $\bar{\boldsymbol{\beta}}$.

Equation (4) indicates that the number of observations on $\boldsymbol{\delta}$ that can be used to estimate $\bar{\boldsymbol{\beta}}$ equals the number of products: effectively we have to estimate $\bar{\boldsymbol{\beta}}$ from the product-level data. Consequently, we cannot identify $\bar{\boldsymbol{\beta}}$ without some assumption on the joint distribution of $(\xi, \mathbf{x})$. This is exactly the same identification problem faced by BLP. As noted in that article and elsewhere (Nevo 2000), different assumptions on the joint distribution of $(\xi, \mathbf{x})$ can be used to identify the remaining parameters. To account for the simultaneity problem, BLP assume that the ξ_j are mean independent of the *nonprice* characteristics of all the products. We make use of this and other possible restrictions below.

To return to the implications of our model, market-level aggregate consumer behavior is obtained by summing the choices implied by the individual utility model over the population's distribution of consumer attributes. Let $\mathbf{w}_i$ be the vector of both the observed $(\mathbf{z}_i)$ and unobserved $(\boldsymbol{\nu}_i, \boldsymbol{\epsilon}_i)$ individual attributes, $\mathbf{w}_i = (\mathbf{z}_i, \boldsymbol{\nu}_i, \boldsymbol{\epsilon}_i)$, and denote its distribution in the population by $\mathcal{P}_{\mathbf{w}}$. The fraction of households that choose good j (aggregate demand) is given by integrating over the set of attributes that imply a preference for good j:

$$s_j(\boldsymbol{\delta}, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u; \mathbf{x}, \mathcal{P}_{\mathbf{w}}) = \int_{A_j(\boldsymbol{\delta}, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u; \mathbf{x})} \mathcal{P}_{\mathbf{w}}(d\mathbf{w}), \tag{5}$$

[4] See also Ichimura and Thompson (1998), who discuss non- and semiparametric identification.

where

$$A_j(\boldsymbol{\delta}, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u; \mathbf{x}) = \{\mathbf{w} : \max_{r=0,1,\ldots,J} [u_{ir}(\mathbf{w}; \boldsymbol{\delta}, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u, \mathbf{x})] = u_{ij}\}.$$

Just as the basic form of equation (1) is familiar from the econometric discrete choice literature (see, e.g., McFadden 1981), the notion of aggregating discrete choices to market demand has been used extensively in the industrial organization literature on product differentiation. An early example is Hotelling (1929); Anderson, DePalma, and Thisse (1992) provide a more recent discussion with extensive references.

III. Estimation

We begin with an outline of our estimation procedure focusing on the role it gives to alternative data sources. The reader who is not interested in the technical detail should be able to proceed directly from subsection *A* to the section that introduces the data (Sec. IV). Subsection *B* explains how we compute the objective function. The Appendix outlines how we construct our standard errors.

A. Outline of the Estimation Procedure

Since our micro data allow us to estimate choice-specific constant terms, we faced a choice of whether to estimate the vector $\boldsymbol{\theta} = (\boldsymbol{\beta}^o, \boldsymbol{\beta}^u, \boldsymbol{\delta})$ or to impose enough additional restrictions on the joint distribution of $(\xi, \mathbf{x})$ to enable us to identify $\bar{\boldsymbol{\beta}}$ and estimate only $(\boldsymbol{\beta}^o, \boldsymbol{\beta}^u, \bar{\boldsymbol{\beta}})$. Formally, the trade-off here is familiar: gaining efficiency from additional restrictions versus losing consistency if those restrictions are wrong.

We chose to estimate $\boldsymbol{\theta}$ without imposing any additional restrictions for two reasons. First, the CAMIP data set is large, so we are not particularly concerned with precision. Second, as noted in BLP, the distribution of $(\xi, \mathbf{x})$ is partly determined by product development decisions, so a priori restrictions on it are hard to evaluate. Our choice implies estimates of $(\boldsymbol{\beta}^o, \boldsymbol{\beta}^u)$ that are robust to assumptions on the $(\xi, \mathbf{x})$ distribution. We then use the estimated $\boldsymbol{\delta}$'s to estimate $\bar{\boldsymbol{\beta}}$ using various assumptions on $(\xi, \mathbf{x})$ (Sec. VI).

Efficiency considerations argue for using maximum likelihood estimates of $\boldsymbol{\theta}$, but this was too computationally burdensome (see app. A of our earlier working paper [Berry et al. 2001]). Therefore, we use a method of moments estimator. This compares the moments predicted by our model for different values of $\boldsymbol{\theta}$ to our sample's moments and then chooses the value of $\boldsymbol{\theta}$ that minimizes the "distance" between the model's predictions and the data.

We matched three "sets" of predicted moments to their data ana-

logues: (1) the covariances of the observed first-choice product characteristics, the $\mathbf{x}$, with the observed consumer attributes, the $\mathbf{z}$ (e.g., the covariance of family size and first-choice vehicle size); (2) the covariances between the first-choice product characteristics and the second-choice product characteristics (e.g., the covariance of the size of the first-choice vehicle with the size of the second-choice vehicle); and (3) the market shares of the J products.

The first set of moments match observed consumer attributes to the characteristics of the chosen vehicles. We think of these moments as particularly useful for estimating β^o, the coefficients on the interactions between observed product characteristics and household attributes ($\mathbf{x}$ and $\mathbf{z}$).[5] If the first-choice car characteristics are denoted by $\mathbf{x}^1$ and $\mathbf{z}$ denotes household attributes, we fit the model's predictions for $E(\mathbf{x}^1\mathbf{z}')$ and for $E(\mathbf{z})$ to their CAMIP sample analogues. We include in $E(\mathbf{x}^1\mathbf{z}')$ a separate moment condition for each interaction term in the utility specification. Since the CAMIP sampling rates are roughly in proportion to market share, the expectation $E(\mathbf{z})$ is roughly the expected value of the attributes of households that chose to buy a car. The $E(\mathbf{z})$ moments are therefore particularly useful in estimating the parameters that define the utility of the outside good.

The second set of moments, between first- and second-choice characteristics, are particularly useful in identifying the importance of the unobserved consumer characteristics. Note that if all relevant consumer attributes were observed ($\beta^u = 0$), then the coefficients of the observed consumer attributes, β^o, would determine both the first- and second-choice vehicle characteristics and hence the correlation between them. If the model with $\beta^u \equiv 0$ predicts a first/second-choice correlation that is much less than the correlation found in the data, we would conclude that the β^u are necessary to explain observed substitution patterns. Our specification has one element of β^u for each included car characteristic, and we include a predicted first/second-choice covariance for each such characteristic.

As noted in Berry (1994), given $\beta \equiv (\beta^o, \beta^u)$, there is a unique δ that matches the observed market shares equal to the model's predicted share. So the third set of moments are particularly useful in estimating the δ parameters.

B. The Fitted Moments

This subsection explains how we compute the moments that go into our method of moments estimation algorithm and considers the limit

[5] If $\beta^o = 0$ and we used only first-choice data, then the aggregate shares used in BLP would be sufficient statistics for the first-choice data, and the match of individuals to the car they chose would contain no additional information.

distribution of the parameter estimates. This requires some additional notation, an introduction to our data sets, and assumptions on the joint distribution of the household attributes.

Let N indicate the number of households in the U.S. population (over 100 million). Then the product-level data consist of J couples, (s_j^N, x_j), where s_j^N is the share of the population that purchased vehicle j, and x_j is a vector of the vehicle's observed characteristics (one of which is price, p_j). The equation $s_0^N = 1 - \sum_j s_j^N$ is the fraction of the population that does not purchase one of our J vehicles. Our model implies that the market share observed in the data, say $\mathbf{s}^N$, distributes multinomially about $\mathbf{s}(\delta_0, \beta_0; \mathbf{x}, \mathcal{P}_w)$, where (β_0, δ_0) represents the true value of that vector, and has a covariance matrix whose elements are all less than N^{-1}.

The consumer-level, or CAMIP, data are a choice-based sample drawn from new vehicle registrations. General Motors determines the number of households to sample from the registrations for each vehicle, say n_j, and then the characteristics of the households sampled and their second-choice vehicles are found. We let $n = \sum_j n_j$ and index the number of households in the CAMIP data by $i = 1, \ldots, n$. The expression $y_i^1 = j$ is our notation for the event that the first choice of household i is vehicle j, and $y_i^2 = k$ indicates that the second choice is vehicle k.

To derive the predictions of the model, we have to specify a joint distribution for the observed and unobserved consumer attributes, the $\mathbf{z}_i$, and the $(\boldsymbol{\nu}_i, \boldsymbol{\epsilon}_i)$ couples. Since the Current Population Survey (CPS) is a random sample of U.S. households, we can use it to sample from $\mathcal{P}_z$ directly. The $(\boldsymbol{\nu}, \boldsymbol{\epsilon})$ couples are assumed to distribute independently of $\mathbf{z}$ and of each other. Recall that the means of these variables go into the constant terms (the $\boldsymbol{\delta}$). We assume that the deviations from the means (our $\boldsymbol{\nu}$) are independent, normal random variables. Thus β_k^u can be interpreted as the standard deviation of the unobserved distribution of tastes for vehicle characteristic k. The sole exception to this is the unobserved characteristic that interacts with price, which is assumed to be lognormal (this allows us to impose the constraint that no one prefers higher prices; see eq. [14] below for more detail). These assumptions give us the marginal distribution of $\boldsymbol{\nu}$, denoted $\mathcal{P}_\nu$.

Finally, for computational simplicity, we assume that the idiosyncratic errors, the ϵ_{ij}, have an independently and identically distributed extreme value "double exponential" distribution. This assumption yields the logit functional form for the model's choice probabilities *conditional* on a $(\mathbf{z}, \boldsymbol{\nu})$ couple:

$$\Pr(y_i^1 = j | \mathbf{z}_i, \boldsymbol{\nu}_i, \boldsymbol{\theta}, \mathbf{x}) = \frac{\exp(\delta_j + \sum_{kr} x_{jk} z_{ir} \beta_{kr}^o + \sum_k x_{jk} \nu_{ik} \beta_k^u)}{1 + \sum_q \exp(\delta_q + \sum_{kr} x_{qk} z_{ir} \beta_{kr}^o + \sum_k x_{qk} \nu_{ik} \beta_k^u)}. \quad (6)$$

Note that the choice probabilities in (6) are an easy to calculate function of $\mathbf{z}$, $\boldsymbol{\nu}$, and $\boldsymbol{\theta}$.

We now move to the computation of our moments. The moments for the aggregate shares are treated slightly differently in order to solve another computational problem. Since we have over 200 car models, $\boldsymbol{\delta}$ has 200 elements, and a search over $\boldsymbol{\theta}$ is a search over about 250 dimensions. Since we cannot search over that many dimensions effectively, we use the aggregate moments to "concentrate out" the $\boldsymbol{\delta}$ parameter and then search only over $\boldsymbol{\beta}$.

Recall that the variance of $\mathbf{s}^N - \mathbf{s}(\delta_0, \beta_0; \mathbf{x}, \mathcal{P}_\mathbf{w})$ is of order N^{-1} and $N^{-1} \approx 0$. Consequently, if we could calculate $\mathbf{s}(\cdot)$ exactly, an efficient method of moments algorithm would chose $\boldsymbol{\theta}$ so that $\mathbf{s}^N \approx \mathbf{s}(\cdot)$. So we (i) use the contraction provided by BLP to find the value of $\boldsymbol{\delta}$ that makes $\mathbf{s}^N \equiv \mathbf{s}(\boldsymbol{\beta}, \boldsymbol{\delta}; \cdot)$, say $\boldsymbol{\delta}(\boldsymbol{\beta}, \mathbf{s}^N; \cdot)$, for each guess at $\boldsymbol{\beta}$; (ii) substitute that $\boldsymbol{\delta}(\boldsymbol{\beta}, \mathbf{s}^N; \cdot)$ for $\boldsymbol{\delta}$ into the model's predictions for the micro moments, making them a function of $(\boldsymbol{\beta}, \boldsymbol{\delta}(\boldsymbol{\beta}, \mathbf{s}^N; \cdot))$; and (iii) then search to find the value of $\boldsymbol{\beta}$ that minimizes the distance between those predictions and the data. This procedure eliminates any need for a search over $\boldsymbol{\delta}$, and the contraction mapping in Berry et al. (1995) solves for $\boldsymbol{\delta}(\boldsymbol{\beta}, \mathbf{s}^N; \cdot)$ quite quickly.

To implement this procedure, we need to compute the market shares predicted by our model for different values of $\boldsymbol{\theta}$, that is, to integrate the probability in equation (6) over the distribution of $(\mathbf{z}, \boldsymbol{\nu})$. Unfortunately, that integral does not have an analytic form. Consequently, we follow Pakes (1986) and use simulation to approximate its value. Specifically, let $(\mathbf{z}_r, \boldsymbol{\nu}_r)$, for $r = 1, \ldots, ns$, index ns random draws on a couple whose first component, $\mathbf{z}_r$, is taken from the CPS and whose second component, $\boldsymbol{\nu}_r$, is taken from the assumed distribution of $\boldsymbol{\nu}$. We then define $\boldsymbol{\delta}^{ns,N}(\boldsymbol{\beta})$ implicitly as the value of this vector that sets[6]

$$G^3_{ns,N}(\boldsymbol{\theta}) = s_j^N - \frac{1}{ns}\sum_{r=1}^{ns} \Pr\left(y^1 = j | \mathbf{z}_r, \boldsymbol{\nu}_r, \boldsymbol{\beta}, \boldsymbol{\delta}^{ns,M}(\boldsymbol{\beta})\right) \tag{7}$$

to zero (and can be found quickly with the BLP contraction mapping).

Note that we draw the $(\mathbf{z}_r, \boldsymbol{\nu}_r)$ couples once at the beginning of the algorithm and hold them constant thereafter. This ensures that the limit theorems in Pakes and Pollard (1989) apply to our estimators. This use of simulation does, however, put simulation error in our estimates of $\boldsymbol{\delta}$ given $\boldsymbol{\beta}$, and this affects the asymptotic variance of the estimates of $\boldsymbol{\beta}$ (see the Appendix).

Next we calculate the model's predictions for the covariances between

[6] In practice we do not just take random draws from the distributions of $\mathbf{z}$ and $\boldsymbol{\nu}$, but rather use importance sampling techniques, analogous to those used in BLP, to reduce the variance of our estimated integrals.

the first-choice car characteristics and household attributes. Since the CAMIP data are choice-based, the moments we have to fit to the data are the model's predictions for the attributes of a household that chose a particular vehicle. To form the sample moment, we interact the average attributes of households that chose vehicle j with the characteristics of that vehicle and then average over the different vehicles (using the CAMIP sampling weights). That is, our first-choice moments are

$$G^1_{n,ns,N}(\boldsymbol{\beta}) \approx \sum_j \frac{n_j}{n} x^1_{kj}\left\{(n_j)^{-1}\sum_{i_j=1}^{n_j} z_{i_j} - E[\mathbf{z}|y^1_i = j, \boldsymbol{\beta}]\right\}, \tag{8}$$

where, at the risk of some misunderstanding, it is now understood that when we condition on $\boldsymbol{\beta}$ we are conditioning on $(\boldsymbol{\beta}, \boldsymbol{\delta}^{ns,N}(\boldsymbol{\beta}; \cdot))$.

We use an approximation sign in equation (8) to indicate that we cannot calculate $E[\mathbf{z}|y^1 = j, \boldsymbol{\beta}]$ exactly. To obtain our approximation we use Bayes' rule to rewrite[7]

$$E[\mathbf{z}|y^1 = j, \boldsymbol{\beta}] = \int_{\mathbf{z}} \mathbf{z}\mathcal{P}(d\mathbf{z}|y^1 = j, \boldsymbol{\beta}) = \frac{\int_{\mathbf{z}} \mathbf{z} \Pr(y^1 = j|\mathbf{z}, \boldsymbol{\beta})\mathcal{P}(d\mathbf{z})}{\Pr(y^1 = j, \boldsymbol{\beta})}$$

and substitute from the model's predictions for the choice probabilities (eq. [6]) to obtain

$$E[\mathbf{z}|y^1 = j, \boldsymbol{\beta}] = \frac{\int_{\mathbf{z}}\int_{\boldsymbol{\nu}} \mathbf{z} \Pr(y^1 = j|\mathbf{z}, \boldsymbol{\nu}, \boldsymbol{\beta})\mathcal{P}(d\mathbf{z}, d\boldsymbol{\nu})}{\Pr(y^1 = j, \boldsymbol{\beta})}. \tag{9}$$

For each value of $\boldsymbol{\beta}$, our model's prediction for the denominator of (9) will, by virtue of the choice of $\boldsymbol{\delta}^{N,ns}(\boldsymbol{\beta})$, exactly equal s^N_j. However, we have to simulate the integral in the numerator. Using the same draws on $(\mathbf{z}_r, \boldsymbol{\nu}_r)$ we used in equation (7), we obtain our approximation as

$$E[\mathbf{z}|y^1 = j, \boldsymbol{\beta}] \approx \frac{(ns)^{-1}\sum_r \mathbf{z}_r \Pr(y^1 = j|\mathbf{z}_r, \boldsymbol{\nu}_r, \boldsymbol{\beta}, \boldsymbol{\delta}^{ns,N}(\boldsymbol{\beta}))}{s^N_J}. \tag{10}$$

The first-choice moments we use are formed by substituting (10) into (8).

An analogous procedure is used to form the moments for the covariances between the characteristics of the first- and second-choice vehicles. Consider only the households whose first choice was vehicle j. For those households, the difference between the average value of characteristic k of the second-choice vehicle they list in their responses and

[7] This follows the literature on choice-based sampling (see Manski and Lerman 1977; Cosslett 1981; Imbens and Lancaster 1994).

the average value of characteristic k for the second-choice vehicles predicted by our model is

$$\left(\frac{1}{n_j}\sum_{i=1}^{n}\sum_{q\neq j} x_{kq}\{y_i^2 = q\}\{y_i^1 = j\}\right) - \left\{E\left[\sum_{q\neq j} x_{kq}\{y_i^2 = q\} \mid y^1 = j, \boldsymbol{\beta}\right]\right\}, \quad (11)$$

where $\{y_i^2 = q\}$ is the indicator function for the event that vehicle q is the second choice. We interact this difference with x_{kj}^1 and use the CAMIP sample weights to average over first choices to obtain the moment

$$G_{n,ns,N}^2(\boldsymbol{\beta}) \approx \sum_j \frac{n_j}{n} x_{kj}^1 \sum_{q\neq j} x_{kq}\left[\left(\frac{1}{n_j}\sum_{i=1}^{n}\{y_i^2 = q\}\{y_i^1 = j\}\right)\right.$$

$$\left. - \int_{\mathbf{z}}\int_{\boldsymbol{\nu}} \Pr(y^2 = q \mid y^1 = j, \mathbf{z}, \boldsymbol{\nu}, \boldsymbol{\beta})\mathcal{P}_{\mathbf{z}}(d\mathbf{z})\mathcal{P}_{\boldsymbol{\nu}}(d\boldsymbol{\nu})\right]. \quad (12)$$

To calculate the expectation in (12), we note that the second-choice probabilities conditional on $(y^1 = j, \mathbf{z}, \boldsymbol{\nu}, \boldsymbol{\beta})$, that is, $\Pr(y^2 = k \mid y^1 = j, \mathbf{z}, \boldsymbol{\nu}, \boldsymbol{\beta})$, are given by the standard "logit" form in (6) modified to take both vehicle j and the outside alternative out of the choice set (this changes the denominator in the choice probability, eliminating both the one and the jth element in the summation sign). After substituting this into the integrand in (12), we approximate that integral by simulation (as in [8]).

We stack $G^1(\cdot)$ and $G^2(\cdot)$ and use the two-step generalized method of moments estimator (see Hansen 1982) of $\boldsymbol{\beta}$ from the stacked moments. Provided that $ns \to \infty$ and $N \to \infty$ as $n \to \infty$, standard arguments show that this estimator is consistent. Since N is large relative to n and ns in our example, we use the limit distribution for $\boldsymbol{\beta}$ that assumes that as $n \to \infty$, $N/n \to \infty$, but ns/n converges to a positive constant (this ensures that we adjust our variances for simulation error). That limit distribution is normal, and the Appendix explains how to obtain consistent estimates of its covariance matrix.

IV. Data

We begin with a description of the CAMIP data. They contain the results of a propriety survey conducted on behalf of the General Motors Corporation (GM) and are generally not available to researchers outside of the company. This survey is a sample from the set of vehicle registrations in the 1993 model year. For each vehicle, a given number of purchasers are sampled. The intent is to create a random sample conditional on purchased vehicle. The sampled vehicles consist of almost

TABLE 1
COMPARISON OF CONSUMER SAMPLES

Income Range	Percentage in CPS	Percentage in CAMIP	CPS Group Mean	CAMIP Mean
0–36,500	64.17	25.00	16.90	25.96
36,500–55,000	16.97	23.16	44.89	45.43
55,000–85,000	12.34	26.71	66.93	67.46
85,000–	6.52	25.13	114.25	148.19
All	100.00	100.00	34.17	72.27
		Other Demographics		
Family size			2.36	2.65
Age of household head			46.80	46.18
Number of kids			.66	.58
Urban			.46	.35
Rural			.25	.35
Suburban			.29	.30

all vehicles sold in the United States in 1993, not just GM products. The subsample we use contains 37,500 observations (see app. C in Berry et al. [2001] for more details).

The CAMIP questionnaire asks about a limited number of household attributes, including income, age of the household head, family size, and place of residence (urban, rural, etc.). We match each of the household attribute questions to a question in the CPS.[8] Table 1 compares the distribution of household characteristics in the CAMIP sample to those in the CPS. Not surprisingly, CAMIP samples disproportionately from higher-income groups. Households that buy new vehicles, especially high-priced ones, tend to have disproportionately high incomes. A more surprising difference between the two samples is that the CAMIP sample is significantly less urban and more rural than the overall U.S. population. Apparently the rural population purchases a disproportionate number of vehicles, which helps explain the high share of trucks in total vehicle sales.

A. The Choice Set

To define a choice set, we need to classify vehicles into a list of distinct models and associate characteristics and quantities sold with those models. Roughly, our list of vehicles was determined by the sampling cells used to form the data GM provided to us (see Berry et al. [2001, app. C] for details). This was detailed enough to allow us to construct a

[8] The match is generally good, although the CPS questions are usually less ambiguously worded than the CAMIP questions. The CAMIP survey does not ask about the education of the household head. There is a question about the education of the driver of the car, but that is hard to match to a question in the CPS.

choice set of 203 vehicles (147 cars, 25 SUVs, 17 vans, and 14 pickup trucks).[9]

The CAMIP survey contains information on the characteristics of the cars actually sold and on their transaction prices (most studies must make do with the characteristics of a "base" model and list prices). As our x_j we used the characteristics of the modal vehicle for each CAMIP vehicle sample cell (i.e., the combination of options that was most commonly purchased), and for our p_j we used the average price of the modal vehicle. Table 2 provides vehicle characteristics by type of vehicle and the definitions of the vehicle characteristics used throughout the paper. There were about 10.6 million vehicles sold in 1993, and they were sold at an average price of $18,500. This gives total sales of about $196 billion. The light truck market alone had sales of $81.2 billion.

Table 3 provides the characteristics of a selected set of vehicles. Many of the interesting implications of our estimates are best evaluated at a vehicle level of aggregation. To give some idea of these implications without overwhelming the reader with details, we display them only for the illustrative sample of 17 vehicles in table 3. These vehicles were selected because they all have sales that are large relative to the sales of vehicles of their type and because, between them, they cover the major types of vehicles sold.[10]

B. Characteristics of the Micro Data

Table 4 provides the mean characteristics of vehicles chosen by the different demographic groups in the CAMIP sample. A number of interactions between observed household attributes and car characteristics stand out including kids with minivan, income with price, rural with pickup and with all-wheel drive, and age and nearly everything.[11] We used this table and others like it to suggest interactions to include in our specification for utility.

One of the very useful features of the CAMIP data is the presence of second-choice information. Table 5 provides information on second choices for our "representative" sample of vehicles. The first column

[9] In most of the runs we used 218 vehicles. However, in the later runs (reported below), we aggregated 15 very expensive vehicles (an average price of $74,000 and a composite market share of 0.3 percent of vehicles sold) into one "super-luxury" model. Because of the very small shares of these luxury cars, this cut computational time considerably without changing the nature of the results.

[10] The list includes 10 cars (three of them luxury cars), a relatively low- and high-priced minivan, a relatively low- and high-priced jeep, a compact and a full-sized pickup, and a full-sized van.

[11] Older households tend to purchase larger (and therefore heavier) cars with both more safety features and more accessories. They also tend to stay away from SUVs and pickups.

TABLE 2
VEHICLE CHARACTERISTICS BY SIZE/TYPE OF VEHICLE

Vehicle Type	Total Q*	Mean Price*	Mean Pass	Mean HP	Mean Safe	Mean Acc	Mean MPG	Mean Allw	Mean PUPayl	Mean SUVPayl	Number of Vehicles
2-passenger car	57.5	28.5	2	7.1	2	4	20	0	0	0	6
4-passenger car	951.3	15.7	4	4.8	1	3	26	.004	0	0	35
5-passenger car	3,829.7	17.5	5	4.7	1	3	23	.005	0	0	84
≥6-passenger car	1,374.1	21.5	6	4.8	1	4	19	0	0	0	22
Minivan	858.3	19.4	7	4.2	1	3	18	0	0	0	13
SUV	1,163.9	23.3	5	4.4	1	3	15	.9	0	1.3	25
Pickup	2,049.2	15.0	3	4.2	1	2	18	.003	2.0	0	14
Van	269.8	25.0	7	4.1	1	3	14	0	0	0	04
Total	10,553.7	18.4	4.9	4.6	1	2.9	20	.11	.39	.14	203

NOTE.—All means are sales weighted. Variable definitions for vehicle characteristics. Q: U.S. sales and leases to consumers (from Polk); price: average price for modal car; HP: horsepower/weight for engine of modal car ("acceleration"); Pass: number of passengers ("size"); MPG: city miles per gallon from Environmental Protection Agency for modal engine/body style; Acc: number of power accessories of modal car (e.g., power windows, power doors); Safe: safety features: sum of antilock brakes plus airbags; Payl: payload in thousands of pounds, for light trucks (from Wards and *Automotive News*); Minivan: dummy equal to one if minivan; SUV: dummy equal to one if SUV; PU: dummy equal to one if pickup; Van: dummy equal to one if full-size van; Sport: dummy equal to one if sports car (as defined by consumer publications); Allw: dummy equal to one if four-wheel or all-wheel drive; PUPayl: PU × Payl; SUVPayl: SUV × Payl.

* In thousands.

TABLE 3
CHARACTERISTICS OF SELECTED VEHICLES

Model	Q*	Price*	Pass	HP	Safe	Acc	MPG	Allw	Miniv	SUV	PU	Van	PUPayl	SUVPayl
Geo Metro	83.7	7.8	4	3.0	0	0	46	0	0	0	0	0	.00	.00
Cavalier	184.8	11.5	5	4.4	1	2	23	0	0	0	0	0	0	0
Escort	207.7	11.5	5	3.6	0	1	25	0	0	0	0	0	0	0
Corolla	140.0	14.5	5	5.0	1	1	26	0	0	0	0	0	0	0
Sentra	134.0	11.8	4	4.7	0	2	29	0	0	0	0	0	0	0
Accord	321.2	17.3	5	4.5	1	4	22	0	0	0	0	0	0	0
Taurus	221.7	17.7	6	4.5	1	4	21	0	0	0	0	0	0	0
Legend	42.5	32.4	5	5.7	2	4	19	0	0	0	0	0	0	0
Seville	33.7	43.8	5	7.9	2	5	16	0	0	0	0	0	0	0
Lexus LS400	21.9	51.3	5	6.5	2	5	18	0	0	0	0	0	0	0
Caravan	216.9	17.6	7	4.3	1	2	19	0	1	0	0	0	0	0
Quest	38.2	20.5	7	3.9	0	4	17	0	1	0	0	0	0	0
Grand Cherokee	160.3	25.9	5	5.4	2	4	15	1	0	1	0	0	0	1.15
Trooper	18.7	22.8	5	4.5	1	4	15	1	0	1	0	0	0	1.21
GMC Full-Size Pickup	141.2	16.8	3	4.2	1	3	17	0	0	0	1	0	2.2	0
Toyota Pickup	175.1	13.8	3	4.4	0	0	23	0	0	0	1	0	1.64	0
Econovan	116.3	24.5	7	3.4	1	3	14	0	0	0	0	1	0	0

* In thousands.

TABLE 4
VEHICLE CHARACTERISTICS OF DIFFERENT DEMOGRAPHIC GROUPS

Group	Price	HP	Pass	Acc	Safe	Sport	MPG	Allw	Miniv	SUV	Van	PUPayl	SUVPayl
Age:													
≤30	16.6	4.7	4.5	2.6	.8	.20	22.0	.13	.03	.15	.001	.24	.18
30–50	20.1	4.8	4.9	3.1	1.1	.15	20.4	.13	.08	.13	.009	.18	.18
>50	22.4	4.9	5.1	3.4	1.3	.07	19.8	.06	.04	.04	.011	.19	.07
Kids:													
0	20.9	4.9	4.8	3.2	1.1	.14	20.4	.10	.03	.09	.006	.20	.12
1	19.2	4.7	4.8	3.0	1.0	.13	21.0	.12	.06	.11	.006	.20	.15
2+	20.1	4.6	5.3	3.1	1.0	.08	19.9	.12	.18	.13	.020	.16	.18
Family size:													
1	19.8	4.9	4.7	3.1	1.1	.20	21.2	.09	.01	.08	.003	.20	.12
2	21.5	4.9	4.9	3.3	1.2	.11	20.1	.10	.04	.09	.007	.20	.12
3+	19.7	4.7	5.0	3.1	1.0	.12	20.5	.11	.10	.12	.012	.19	.16
Urban	20.6	4.8	4.9	3.2	1.1	.13	20.7	.10	.05	.10	.009	.14	.14
Suburban	21.7	5.0	4.9	3.4	1.2	.15	20.3	.10	.06	.10	.006	.10	.14
Rural	19.2	4.7	4.9	3.0	1.0	.11	20.2	.12	.06	.11	.010	.31	.14
Income:													
≤37,000	16.6	4.6	4.8	2.6	.88	.12	21.9	.08	.04	.07	.008	.25	.08
37,000–55,000	18.5	4.7	4.9	3.0	1.0	.12	20.7	.10	.07	.10	.011	.24	.13
55,000–85,000	20.3	4.8	4.9	3.2	1.1	.14	20.0	.13	.07	.13	.009	.19	.17
>85,000	26.3	5.2	4.9	3.7	1.4	.14	19.1	.11	.05	.12	.006	.08	.17

TABLE 5
EXAMPLES OF SECOND CHOICES

Model	n_j	Modal Second Choice	Number Choosing	Next Second Choice	(Modal+ Next)/n	Number of Different Choices
Metro	188	Escort	22	Geo Storm	.22	49
Cavalier	238	Escort	16	LeBaron	.12	59
Escort	166	Tempo	16	Taurus	.18	53
Corolla	250	Civic	42	Camry	.33	55
Sentra	203	Corolla	34	Civic	.31	60
Accord	223	Camry	58	Taurus	.35	61
Taurus	147	Camry	18	Sable	.22	45
Legend	119	Lexus ES300	19	Lexus SC300	.24	40
Seville	243	DeVille	38	Lincoln MK8	.26	49
Lexus LS400	148	DeVille	33	Infiniti Q45	.39	27
Caravan	166	Voyager	31	Aerostar	.32	36
Quest	232	Caravan	50	Villager	.43	31
Grand Cherokee	137	Explorer	75	Blazer	.59	34
Trooper	137	Explorer	43	Rodeo	.41	27
GMC Full-Size Pickup	469	Chevy Full-Size Pickup	222	Ford Full-Size Pickup	.55	29
Toyota Pickup	113	Ford Ranger	29	Nissan Pickup	.43	25
Econovan	90	Chevy Full-Size Van	20	Suburban	.44	23

gives the first-choice vehicle, and the second column gives the CAMIP sample size n. The next columns, in order, give the modal second choice, the number of sampled consumers making that choice, the second choice with the second-highest number of consumers, the fraction of n that chose one of the two second choices listed, and the number of different second choices made. For example, the sample contains 166 purchasers of the Ford Escort. Their modal second choice was the Ford Tempo, whereas the second choice with the next-highest number of consumers was the Ford Taurus. Together these two second choices accounted for 39, or 18 percent, of the consumers who chose the Escort. There were 51 other second choices registered among Escort purchasers.

There are a large number of different second choices for the same first-choice car, but the second choices are more concentrated for light trucks and for higher-priced cars. Note also that the second choice is often produced by the same company as the first-choice car, a fact that argues strongly for pricing policies that maximize the joint profits of the firm across all the products it produces.

As expected, the second-choice vehicles have characteristics that are similar to those of the first choices. The correlations of the different vehicle characteristics across the first and second choices of the households were all positive and highly significant (the correlations for price and minivan were largest, about .7; those for MPG, size, and other type dummies were about .6; and the rest were between .3 and .5). Unfortunately, the surveyed consumers are not asked whether they would have purchased a vehicle at all if their first choice had not been available, so we cannot provide any descriptive evidence on how many consumers might substitute out of the new vehicle market altogether if their first choice were unavailable.[12]

V. The Estimates of β^o and β^u

We begin with details of our specification. Recall that utility (eq. [1]) has interaction terms of the form $\sum_k \tilde{\beta}_{ik} x_{jk}$, where k indexes characteristics, i indexes households, and j indexes products. For all characteristics except price, we assume that

$$\tilde{\beta}_{ik} = \bar{\beta}_k + \sum_r z_{ir}\beta^o_{kr} + \beta^u_k \nu_{ik}. \tag{13}$$

As in (2), the $\bar{\beta}$'s are subsumed in the product-specific constants, δ, and

[12] Some households listed a second choice that was broader than our first-choice cells (e.g., a Ford pickup). The empirical analysis explicitly aggregates the respective cell probabilities for the second choices of these consumers.

the ν's are assumed to have independent (across both consumers and characteristics) standard normal distributions. Thus the β^u are the standard deviations of the contribution of unmeasured consumer attributes to the variance in the marginal utility for characteristics k. We let the descriptive tables and a number of preliminary runs guide our choice of which $\mathbf{z}_i$ to interact with the different x_j. Observed interactions were dropped from our early runs if we found them to be consistently unimportant.[13]

We assume the price coefficient to be a function of effective wealth, say W, and then model W in terms of household attributes. That is, our price coefficient is $-e^{-W}$, so that its log is a decreasing function of

$$W_i \equiv \sum_r z_{ir}\beta^o_{w,r} + \beta^u_w \nu_{iw}. \tag{14}$$

Initially the $z_{i,r}$ included a constant, family size, a spline in income that was allowed to change derivatives at each of the quartiles of the CAMIP income distribution, and a lognormally distributed $\nu_{i,w}$ (for determinants of wealth not contained in our data). The data indicated needed only a change in the derivative of the income/price interaction in the spline at the seventy-fifth income percentile.

We have little a priori information on the outside option of not buying a car, so in early runs we let it be a linear function of all observed household attributes, a random normal disturbance, and the "logit" error. These runs indicated that the only attributes that mattered were income, family size, and, sometimes, the number of adults.

Tables 6 and 7 provide the estimates from our full model (col. 1 in each table) and compare them to those from more traditional models. Table 6 presents estimates of the β^o coefficients of interactions with observed household attributes, and table 7 presents estimates of the β^u coefficients of interactions with unobserved attributes. There are three comparison models. The first two are obtained from our full specification but with $\beta^u = 0$, giving us a standard logit model with closed-form probabilities. This model has *both* choice-specific intercepts and interactions between *observed* household attributes and vehicle characteristics (so we still have to use simulation to obtain predictions for aggregate shares; see also Berry et al. [2001, app. A]). Column 2 of table 6 provides the estimates obtained when using only first-choice data, and column 3 provides the estimates using both first- and second-choice data. The third comparison model sets $\beta^o = 0$, and so does not appear in table 6 (just in table 7). This model is like the BLP model in that it has no observed consumer attributes.

[13] Our use of preliminary runs gives us some confidence that our results are reasonably robust to the inclusion of further interactions. However, it makes our standard errors suspect in the usual way.

TABLE 6
ESTIMATES OF INTERACTION TERMS, β^o

VEHICLE CHARACTERISTIC AND HOUSEHOLD ATTRIBUTE	FULL MODEL (1)	LOGIT	
		First (2)	First and Second (3)
Price:			
Constant	−2.18	.092	.139
	(.142)	(.0001)	(.0003)
Income × (income <75th percentile)	.714	.299	.344
	(.044)	(.002)	(.001)
Income × (income >75th percentile)	1.17	.466	.603
	(.083)	(.091)	(.007)
Family size	−.565	−.144	−.143
	(.010)	(.001)	(.006)
Minivan: Kids (kids have age ≤16)	1.973	.765	.771
	(.242)	(.098)	(.323)
Pass:			
Adults (adults have age >16)	.203	.018	−.067
	(.095)	(.0004)	(.009)
Family size	.536	−.055	−.006
	(.052)	(.003)	(.0002)
Age (of household head)	.019	.002	.005
	(.003)	(.00001)	(.00001)
HP: Age	−.002	−.010	−.012
	(.001)	(.0004)	(.0001)
Acc:			
Age	.0004	.001	−.002
	(.001)	(.00001)	(.0001)
Age^2	.0001	.000	.000
	(.00001)	(.00001)	(.00001)
PUPayl:			
Age	.0174	−.003	.000
	(.002)	(.0001)	(.00001)
Rural dummy	1.075	.512	.376
	(.179)	(.005)	(.008)
Safe: Age	.013	.015	.016
	(.0006)	(.001)	(.0004)
SUV:			
Age	−.219	−.043	−.043
	(.010)	(.003)	(.004)
Rural dummy	.332	.403	−.016
	(.156)	(.007)	(.002)
Allw: Rural dummy	.278	.142	.734
	(.247)	(.005)	(.246)
Outside good:			
Total income	5.151	−.228	−.305
	(.228)	(.096)	(.063)
Family size	−.007	.532	−.346
	(.002)	(.057)	(.004)
Adults	−.428	.851	1.953
	(.766)	(.112)	(.148)

TABLE 7
ESTIMATES OF INTERACTION TERMS, β^u

Parameter Name	Full Model (1)	$\beta^o \equiv 0$ (2)
Price	.449	.055
	(.026)	(.004)
HP	.030	.183
	(.016)	(.020)
Pass	2.74	1.444
	(.147)	(.055)
Sport	.002	2.763
	(.0004)	(.068)
Acc	.554	.515
	(.078)	(.055)
Safe	.260	.376
	(.130)	(.093)
MPG	.488	.430
	(.018)	(.017)
Allw	.740	.431
	(.179)	(.049)
Minivan	4.787	6.641
	(.353)	(.113)
SUV	3.076	3.231
	(.292)	(.114)
Van	1.713	6.888
	(.289)	(.266)
PUPayl	2.160	4.301
	(.092)	(.210)
SUVPayl	.356	.015
	(.072)	(.013)
Chrysler	1.689	1.383
	(.058)	(.051)
Ford	.915	1.410
	(.072)	(.051)
GM	1.885	1.844
	(.057)	(.105)
Honda	.329	.086
	(.128)	(.043)
Nissan	.506	1.588
	(.142)	(.071)
Toyota	.169	.576
	(.134)	(.094)
Small Asian*	1.467	2.155
	(.068)	(.022)
European*	.454	1.883
	(.084)	(.034)
Outside good	27.858	10.256
	(1.004)	(.506)

* We constrained the coefficients on the dummies for the different European firms to be the same, and we did the same for the smaller Asian producers.

There was one other comparison model we tried to estimate: our full model using only the first-choice data (like the results in col. 2). However, even after *substantial* experimentation, we had convergence problems with these runs, and it eventually became clear that very different parameter values could generate values of the objective function that were essentially the same as that of the minimum of that function. Apparently it is the availability of second-choice data that enables us to focus in on a set of precise parameter estimates. Note that since we have only a single cross section, there is no variance in the choice set across observations.[14] In applications to other data sets, variation in the choice set (either over time or across markets) might provide the information necessary to estimate the random coefficients.

The first four rows of table 6 show that all three observed interactions with price are sharply estimated and have the expected sign (all else equal, larger families have lower "wealth"). Indeed almost all interactions in table 6 both had an expected sign and were precisely estimated in *all three* specifications.[15] In addition to the price interactions, this includes the interactions between minivans and kids (+), age and passengers (+), age and safety (+), HP and age (−), SUV and age (−), and rural and pickup payload (+).

The full model had only one parameter estimate that might be considered an anomaly (the positive age/pickup payload interaction); the first-choice logit estimates had as their sole clear anomaly a negative interaction between number of passengers and family size (and the implication of this is ameliorated by the highly positive interactions between the minivan dummy and kids and between adults and passenger size). The second-choice logits do a little worse, predicting negative interactions between family size and passengers and between rural and the SUV dummy. The logits also have a pattern of outside good coefficients that is counterintuitive. While estimates from our full model imply that households with more income and smaller families tend to have larger values for the outside option, the logits predict the opposite.[16] However, the outside good's coefficients are reduced form and hence are more difficult to interpret.

On the whole the logits performed quite well in terms of producing

[14] A referee noted that random coefficients models have been found unstable in many related cross-sectional contexts. For a review of random coefficients models, see McCulloch, Polson, and Rossi (2000) and the literature cited there.

[15] We did not present the breakdown of the variance in the estimated coefficients into portions caused by simulation and sampling error, but typically somewhat less than half of this variance is due to simulation.

[16] Note that though our full model predicts a higher value of the outside good for higher-income people, it also predicts a higher probability of purchasing a vehicle for higher-income people, since the negative price interactions with income more than offset the positive interactions with the outside good.

sensible signs for coefficients, so the increased computational burden of the full model is not obviously justified by the pattern of estimated interactions between **x** and **z**. However, while the demographic interaction terms both seem to make sense and are sharply estimated, table 7 indicates that they apparently *do not* explain the full pattern of substitution in the data. The estimated β^u coefficients are large and very precisely estimated. No matter how many observed interactions we allowed for, we needed numerous additional unobserved interactions to explain the data. Of course if we had richer consumer data, we would hope to capture more with household observables; but the CAMIP data do have most of the household attributes generally available in large consumer choice data sets.

As shown in table 7, 19 out of 22 coefficients are highly significant (11 with *t*-values over 10) and two are marginally significant. Interestingly, there seems to be a wider dispersion of preferences for vehicles of U.S. companies than for those of Japanese companies. The model with no observed attributes has even more precisely estimated β^u coefficients (col. 2) since it has fewer other coefficients to estimate. Indeed the $\beta^o \equiv 0$ model has *all* β^u coefficients significant and several with *t*-values over 50.

A clear pattern emerged when we compared the fit of the various models. The full model fit the (uncentered) moments derived from the interactions between observed consumer attributes and first-choice car characteristics (eq. [8]) about as well as the first- and the second-choice logits did, whereas the model with no observed interactions could not fit these moments at all. On the other hand, the model with no observed interactions fit the (uncentered) covariance of the first- and second-choice car characteristics (eq. [12]) about as well as the full model did, but the percentage errors in the first- and second-choice logits for these moments were typically five to 10 times as large.

The logits, then, provide an adequate fit for the correlations between observed household and vehicle characteristics but do very poorly in matching the characteristics of the first- and second-choice car. This might lead us to believe that the logits will predict the demographics of consumers well but do a poor job of predicting substitution patterns. The model with no observed attributes provides an adequate fit for the correlations of the characteristics of the first- and second-choice car but has no prediction at all for the correlations between the observed household and the observed vehicle characteristics. Our full model (which nests all specifications) does about as well as the best of the alternatives in both these dimensions.

VI. $\bar{\beta}$ and Substitution Patterns

The only demand parameters left to estimate are the $\bar{\boldsymbol{\beta}}$, the effects of the characteristics on the choice-specific intercepts (the $\{\delta_j\}$). Recall that

$$\delta_j = p_j\bar{\beta}_p + \sum_{k \neq p}^{K} x_{jk}\bar{\beta}_k + \xi_j. \tag{15}$$

The problems encountered in estimating equation (15) are similar to the problems discussed in BLP in the context of estimating demand systems from product-level data. In particular, consistent estimation of (15) requires instruments at least for the endogenous prices. Note that in contrast to our single 1993 cross section, BLP had 20 annual cross sections. Still their estimates that used only the demand system were too imprecise to be useful. This suggests that we also will have a precision problem, but this time for only a subset of the parameters, $\bar{\boldsymbol{\beta}}$.

A number of additional sources of information could be used to increase the precision of the estimated $\bar{\boldsymbol{\beta}}$. First, we could mimic the BLP study. The authors assumed (i) a functional form for marginal costs and (ii) a Nash equilibrium in prices. This generates a pricing equation that can be used in conjunction with the $\boldsymbol{\delta}$ equation to increase the precision of our estimates of $\bar{\boldsymbol{\beta}}$. In particular, if marginal costs are given by

$$mc_j = \sum_k x_{kj}\gamma_k + \omega_j, \tag{16}$$

where ω_j is an unobserved productivity term that is mean independent of $\mathbf{x}$ and the γ are a set of parameters to be estimated, then the equilibrium assumption implies that price is equal to marginal cost plus a markup:

$$p_j = \sum x_{kj}\gamma_k + b(\mathbf{x}, p, \boldsymbol{\delta}, \bar{\beta}_1, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u)_j + \omega_j, \tag{17}$$

where the form of $b(\mathbf{x}, p, \boldsymbol{\delta}, \bar{\beta}_1, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u)$ is determined by the demand-side parameters and the Nash pricing assumption.

With single-product firms, the markup would be the (familiar) inverse of the semi-elasticity of demand with respect to price. Since we have multiproduct firms, we must use the more complex formula for that case (see, e.g., Berry et al. 1995).

The equilibrium markup in (17) is determined, in part, by ξ, ω, and p and hence needs to be instrumented when that equation is estimated. In addition to x_j, the instruments we use are predictions of the markup:

$$\hat{b}_j \equiv b_j(\mathbf{x}, \hat{p}, \hat{\boldsymbol{\delta}}, \hat{\bar{\beta}}_1, \hat{\boldsymbol{\beta}}^o, \hat{\boldsymbol{\beta}}^u)_j, \tag{18}$$

where $(\hat{\boldsymbol{\delta}}, \hat{p})$ are obtained by projecting our estimate of $\boldsymbol{\delta}$ and the observed p onto the $\mathbf{x}$'s, and $\hat{\bar{\beta}}_p$ is obtained from an initial instrumen-

tal variable estimate of the δ equation. So $\hat{b}_j$ is a function only of the $\mathbf{x}$'s and consistent parameter estimates.[17]

Notice that this method of identifying $\bar{\boldsymbol{\beta}}$ relies on our pricing assumption (though our estimates of $(\boldsymbol{\beta}^o, \boldsymbol{\beta}^u)$ do not) and relies quite heavily on functional form restrictions (we do not observe multiple prices for a given vehicle). This suggests looking for other ways of identifying $\bar{\boldsymbol{\beta}}$. Moreover, since the equilibrium markups and price elasticities depend only on the coefficients estimated in the first-stage analysis and on $\partial\delta_j/\partial p_j$ and equation (15) implies that $\partial\delta_j/\partial p_j = \bar{\beta}_p$, we can analyze all price change effects from the estimates of $(\boldsymbol{\delta}, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u)$ and any single restriction that identifies $\bar{\beta}_p$.[18] On the basis of their experience, the staff at General Motors suggested that the aggregate (market) price elasticity in the market for new vehicles was near one. An alternative estimate of $\bar{\beta}_p$ is then the value that sets the 1993 market elasticity equal to one.

When we use the δ equation (15) alone, the instrumental variable estimates of $\bar{\boldsymbol{\beta}}$ are too imprecise to be of much use (our estimate of $\bar{\beta}_p$ had a standard error 10 times the point estimate: 25 vs. 2.5). The instrumental variable estimate of $\bar{\beta}_p$ from the two-equation model (which uses the δ equation *and* the pricing assumption) is -3.58 and has a standard error of 0.22. The estimate of $\bar{\beta}_p$ that "calibrates" to GM's market elasticity of -1 is -11. We consider these two estimates as well as the estimate implicit in studies that ignore the correlation between the product-specific constant terms and price: $\bar{\beta}_p = 0$.

Table 8 examines the implications of these three estimates of $\bar{\beta}_p$. Panel A provides the implied average (across vehicles) price semi-elasticities and total market price elasticities. Panel B presents the coefficients obtained from the projection of the implied price semi-elasticities onto car characteristics.

Clearly the *level* of the price elasticities increases with the value of the estimate of $\bar{\beta}_p$. On the other hand, the *pattern* of the elasticities seems fairly robust across our estimates of $\bar{\beta}_p$ and accords well with industry reports (especially to reports circa 1993). Semi-elasticities decrease in price, and given price, vans (both mini and full-sized), pickups, SUVs, and, to a lesser extent, sports cars have noticeably smaller elasticities than other vehicles. This goes a long way toward explaining reports of high markups for these vehicles.

We now come to the patterns of substitution across cars. The two types

[17] Actually we iterate on this procedure several times; i.e., we use an initial simple instrumental variable estimate from the δ equation alone to produce our first estimate of $\hat{b}$. Then we construct $\hat{b}$ and use it in a method of moments routine based on the orthogonality conditions from both equations. This produces a new estimate for $\bar{\beta}_p$, which is used to produce another estimate of $\hat{b}$, which was used in another method of moments routine. We continued in this way until convergence.

[18] Similarly, if we were interested in elasticities with respect to any other characteristic, say MPG or HP, we would require only the $\bar{\boldsymbol{\beta}}$ associated with the characteristic of interest.

TABLE 8
IMPLICATIONS OF ALTERNATIVE ESTIMATES OF $\bar{\beta}_p$

	VALUE OF $\bar{\beta}_p$		
	0	−3.58	−11
	A. Implied Average across Vehicles		
Mean semi-elasticity	−.75	−3.94	−10.56
Total market elasticity	−.2	−.4	−1
	B. Coefficients from Projecting Semi-elasticities		
Price	−.016	−.031	−.063
	(.003)	(.006)	(.014)
HP	.023	−.025	−.122
	(.025)	(.044)	(.102)
Pass	.023	.057	.127
	(.029)	(.052)	(.121)
Sport	−.235	−.230	−.219
	(.069)	(.117)	(.273)
Acc	−.086	−.066	−.023
	(.023)	(.040)	(.093)
Safe	−.177	−.137	−.052
	(.038)	(.067)	(.126)
MPG	.010	−.034	−.126
	(.007)	(.013)	(.029)
Allw	.084	.275	.671
	(.103)	(.182)	(.425)
Minivan	−.174	−.730	−1.882
	(.099)	(.174)	(.406)
SUV	−.480	−.923	−1.841
	(.179)	(.316)	(.735)
Van	−.339	−1.112	−2.714
	(.154)	(.272)	(.633)
PUPayl	−.173	−.625	−1.562
	(.050)	(.088)	(.204)
SUVPayl	−.107	−.058	−.400
	(.101)	(.144)	(.416)

NOTE.—Firm dummies are suppressed.

of substitution patterns we consider are (i) substitution induced by price changes and (ii) substitution induced by deleting vehicles from the choice set. The two sets of substitution patterns differ because when price increases, only a selected sample of consumers who purchased the given vehicle substitute out of that vehicle (the more price-sensitive consumers), whereas when a vehicle is deleted from the choice set, all of them must make an alternative choice. These substitution patterns were virtually independent of the estimates of $\bar{\beta}_p$, so we present only one set of results (with $\bar{\beta}_p = -3.58$).

Table 9 presents our model's predictions for the substitution patterns that would result from a small increase in price of the vehicle in the first column. The table provides the name of the vehicle chosen by the

TABLE 9
PRICE SUBSTITUTES FOR SELECTED VEHICLES: ESTIMATES FROM THE FULL MODEL

Vehicle	Price	Semi-elasticity	Best Substitute	Price	Movers* (%)	Second Best	Price	Movers* (%)	To Outside† (%)
Metro	7.84	−1.77	Tercel	9.70	14.96	Festiva	7.41	10.57	17.96
Cavalier	11.46	−4.08	Escort	11.49	8.62	Tempo	10.78	6.80	6.81
Escort	11.49	−4.02	Tempo	10.78	8.21	Cavalier	11.49	7.29	6.56
Corolla	14.51	−3.92	Civic	14.00	8.08	Escort	11.49	7.91	5.00
Sentra	11.78	−3.79	Civic	14.00	13.36	Escort	11.49	4.70	6.55
Accord	17.25	−3.92	Camry	18.20	8.60	Civic	13.00	4.47	5.06
Taurus	17.65	−3.73	Accord	17.25	6.25	Mercury Sable	18.66	6.09	3.97
Legend	32.42	−3.73	Accord	17.25	3.96	Camry	18.20	3.87	4.38
Seville	43.83	−3.16	DeVille	34.40	10.12	El Dorado	35.74	8.04	5.57
Lexus LS400	51.29	−3.43	Mercedes 300	47.71	7.97	Lincoln Town Car	35.68	6.29	5.87
Caravan	17.56	−3.32	Voyager	17.59	35.11	Aerostar	18.13	10.19	5.20
Quest	20.55	−3.98	Aerostar	18.13	12.50	Caravan	17.56	10.38	5.48
Grand Cherokee	25.84	−3.06	Explorer	24.27	17.60	Cherokee	20.10	9.51	6.38
Trooper	22.78	−3.96	Explorer	24.27	17.53	Grand Cherokee	25.85	8.50	5.42
GMC Full-Size Pickup	16.76	−3.78	Chevy Full-Size Pickup	16.78	43.74	Ford Full-Size Pickup	16.68	13.56	6.03
Toyota Pickup	13.77	−3.34	Ranger	11.74	20.53	Nissan Pickup	11.10	11.93	9.35
Econovan	24.54	−2.86	Chevy Van	25.96	12.90	Dodge Van	23.71	9.73	5.38

* Of those who substitute away from the given good in response to the price change, the fraction who substitute to this good.
† Of those who substitute away from the given good in response to the price change, the fraction who substitute to the outside good.

largest fraction of the substituting consumers, the price of that vehicle, and the fraction of those who substitute out of the first-choice vehicle who move to that "best" substitute. It then provides the same information for the vehicle chosen by the second-highest fraction of the substituting consumers. The last column of the table provides the fraction of the substituting consumers who substitute to the outside alternative. Thus the best (price) substitute for the Toyota Corolla is the Honda Civic and the second-best is the Ford Escort. Together these two cars account for about 25 percent of those who substitute out of the Corolla when its price rises. About 5 percent of those who substitute out do not purchase a car at all.

The substitution patterns in table 9 make a lot of sense. Both substitutes tend to be the same type of vehicle as the vehicle whose price rose (minivans substitute to minivans etc.). Among vehicles of the same type, the substitutes tend to be vehicles with prices and size similar to those of the car whose price increased.

Table 10 compares best price substitutes from our model to those from our comparison models. It is clear that the intuitive features of the predictions of our model *are not* shared by the results from the logit models but are, for the most part, shared by the results from the model with no observed attributes. The first-choice logit predicts the Dodge Caravan, a minivan, to be the "best substitute" for nine of the 10 first-choice cars and predicts the Ford Econovan to be the best substitute for the tenth car (a 400 series, or "high-end," Lexus). It also predicts the Dodge Caravan to be the best substitute for both pickups, both SUVs, and the full-size van. The first- and second-choice logit has the Ford full-sized pickup as the best substitute for *all 10* cars.

Apparently the observed characteristics of households do not capture enough of the variation in individual tastes to produce reasonable substitution patterns.[19] On the other hand, the model with no observed attributes ($\beta^o \equiv 0$) produces the same best substitutes as our full model in 12 out of the 17 cases (though its substitutes for the Escort and, to a lesser extent, for the Metro seem questionable). If we are primarily interested in substitution patterns, allowing for interactions between unobserved consumer and product characteristics seems far more important than allowing for the interactions between the observed consumer and product characteristics in our data. Again, recall that our consumer-level data contain most of the variables that are generally available in large micro data sets.

Because of our second-choice data, we are able to compare the mod-

[19] This might have been expected from the logit's inability to fit the moments for the characteristics of the first- and second-choice cars. Note that this is in spite of our allowing for choice-specific constant terms.

TABLE 10
PRICE SUBSTITUTES FOR SELECTED VEHICLES: A COMPARISON AMONG MODELS

		LOGIT		
VEHICLE	FULL MODEL	First	First and Second	SIGMA ONLY
Metro	Tercel	Caravan	Ford Full-Size Pickup	Civic
Cavalier	Escort	Caravan	Ford Full-Size Pickup	Escort
Escort	Tempo	Caravan	Ford Full-Size Pickup	Ranger
Corolla	Escort	Caravan	Ford Full-Size Pickup	Civic
Sentra	Civic	Caravan	Ford Full-Size Pickup	Civic
Accord	Camry	Caravan	Ford Full-Size Pickup	Camry
Taurus	Accord	Caravan	Ford Full-Size Pickup	Accord
Legend	Town Car	Caravan	Ford Full-Size Pickup	Town Car
Seville	DeVille	Caravan	Ford Full-Size Pickup	DeVille
Lexus LS400	Mercedes 300	Econovan	Ford Full-Size Pickup	Seville
Caravan	Voyager	Voyager	Voyager	Voyager
Quest	Aerostar	Caravan	Caravan	Aerostar
Grand Cherokee	Explorer	Caravan	Chevy Full-Size Pickup	Explorer
Trooper	Explorer	Caravan	Chevy Full-Size Pickup	Rodeo
GMC Full-Size Pickup	Chevy Full-Size Pickup	Caravan	Chevy Full-Size Pickup	Chevy Full-Size Pickup
Toyota Pickup	Ranger	Caravan	Chevy Full-Size Pickup	Ranger
Econovan	Dodge Van	Caravan	Ford Full-Size Pickup	Dodge Van

els' predictions for substitution patterns to the data. Table 11 provides the most popular second choice as predicted by the four models. These are the "best substitutes" when the good in the first column is taken off the market. We also ranked the actual data on second choices and placed the data rank of the model's best substitute next to the name of the predicted substitute. Thus, if the Honda Accord were taken off the market, both our model and the $\beta^o = 0$ model predict that the biggest beneficiary would be the Toyota Camry; the data indicate that the Camry is in fact the most popular second choice among Accord purchasers. Our full model predicts exactly the same best substitute as the data nine out of 17 times, predicts one of the top three best substitutes 15 out of 17 times, and never picks a best substitute that the data rank higher than tenth (out of over 200 possible models). The model with $\beta^o \equiv 0$ predicts the same best substitute as the data 12 out of 17 times but has two best substitutes that the data rank above 10.[20] Meanwhile, the logit models (i.e., $\beta^u \equiv 0$) perform as poorly here as they did in table 10, with the Ford full-size pickup being predicted as the best substitute for every car in all the logit specifications. Note also that the best price substitutes and the best second choices are different for about half the cars and one of the light trucks.

VII. Prediction Exercises

Having shown that the implications of our estimate are consistent with available information, we move on to two prediction exercises. First, we evaluate the potential demand for new models; in particular we introduce "high-end" SUVs. Second, we use the system to evaluate a major production decision: shutting down the Oldsmobile division of General Motors. We ask what Oldsmobile purchasers would do were the cars they bought not available. These examples were chosen for their relevance. Several new SUVs were introduced in the late 1990s (an apparent response to the high markups being earned on those vehicles in the period of our data; see table 8), and GM announced its intention to close down its Oldsmobile division in 2000.

Two caveats are worth noting before we go to the results. First, all the data used in our investigations are 1993 data. The market has changed since 1993, and those changes might well affect our estimates. Second, in the exercises done here, we do not allow other actors in the market to respond to the change we are investigating. That is, when

[20] The one set of substitutes that might be considered an anomaly are the predicted substitutes for the Legend. Our model predicts the much cheaper Civic, which is in fact the choice of a small though significant number of Legend buyers. The $\beta^o = 0$ model predicts the Lincoln Town Car, which is priced close to the Legend, but in fact Legend consumers almost never indicate it as a second choice.

TABLE 11
MOST POPULAR SECOND CHOICES: A COMPARISON AMONG MODELS AND TO THE DATA

Vehicle	Full Model	Rank	Logit First	Rank	Logit First and Second	Rank	$\beta^o \equiv 0$	Rank
Metro	Chevy Geo Storm	2	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Tercel	12
Cavalier	Sun Bird	3	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Ford Escort	1
Escort	Tempo	1	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Tempo	1
Corolla	Escort	6	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Civic	1
Sentra	Civic	2	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Civic	2
Accord	Camry	1	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Camry	1
Taurus	Mercury Sable	2	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Accord	4
Legend	Civic	10	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Town Car	≥25
Seville	Deville	1	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	Deville	1
Lexus LS400	MB 300	3	Ford Full-Size Pickup	≥25	Ford Full-Size Pickup	≥25	DeVille 2	1
Caravan	Voyager	1	Ford Full-Size Pickup	≥25	Voyager	1	Voyager	1
Quest	Aerostar	7	Ford Full-Size Pickup	≥25	Caravan	1	Caravan	1
Grand Cherokee	Explorer	1	Chevy Full-Size Pickup	≥25	Chevy Full-Size Pickup	≥25	Explorer	1
Trooper	Explorer	1	Chevy Full-Size Pickup	22	Chevy Full-Size Pickup	22	Rodeo	2
GMC Full-Size Pickup	Chevy Full-Size Pickup	1	Chevy Full-Size Pickup	1	Ford Full-Size Pickup	2	Chevy Full-Size Pickup	1
Toyota Pickup	Ranger	1	Chevy Full-Size Pickup	4	Chevy Full-Size Pickup	4	Ranger	1
Econovan	Chevy Van	1	Ford Full-Size Pickup	6	Ford Full-Size Pickup	6	Chevy Van	1

the Oldsmobile division is shut down, we do not allow for either a realignment of the prices of other products in response to the shutdown or the introduction of the new models that might follow such a shutdown. Similarly, when a new model is introduced, we investigate demand responses under the twin assumptions that prices of other vehicles do not respond to the introduction of that model and that no further new vehicles are introduced.

It is not much more difficult to modify our procedure to find a set of prices that would be a Nash equilibrium to the situation we study. This would, however, require (i) estimates of costs as well as of demand functions and (ii) an assumption on how prices are set. In the past when we have tried similar exercises, we found the impact of the price response to be "second"-order in cases similar to the cases we investigate here but to be central to the analysis of other issues.[21] On the other hand, we have done very little that examines the longer-term responses of the other characteristics (other than price) of the vehicles marketed to changes in the environment.

A. *New Models*

The two new models introduced into the 1993 market are a new Mercedes and a new Toyota SUV. Both new models were introduced with all characteristics *but* price and the unobserved characteristic (i.e., ξ) set equal to the characteristics of the Ford Explorer. The Explorer was the biggest-selling SUV in 1993.

Recall that ξ captures the effect of all the detailed characteristics that are omitted from our specification; we think of it as an "unobserved quality." The ξ of the new Toyota SUV was set equal to the mean ξ of all Toyota cars marketed in that year, and the price of that vehicle was obtained from a regression of price onto a large set of vehicle characteristics and company dummies. This latter regression had a very good fit, and using it allowed us to avoid using the explicit pricing and cost assumptions that would be needed to obtain price from a more complete model. The ξ and p of the new Mercedes SUV were set in the same way using the "low end" of the Mercedes vehicles marketed in 1993.[22] Both

[21] These studies used product-level data and the BLP methodology. Induced price effects were second-order in our analysis of the response of demand to the increase in gas prices in the early 1970s (Pakes et al. 1993). However, we found the price effects to be central in our analysis of voluntary export restraints (Berry et al. 1999) and in an unpublished analysis of particular mergers.

[22] The mean quality and price of the Mercedes were much higher than the quality and price of any SUV marketed at the time. So if we used the means of the Mercedes, we would have been making predictions way out of the range of the data we used in our estimation (and probably also out of the range of the SUV eventually marketed by Mercedes).

TABLE 12
INTRODUCING A MERCEDES SUV

Model	Price	Old Share	New Share	New − Old Share
New car	33,659	.0000	.0762	.0762
		Biggest Declines in Sales		
Ford Explorer	24,274	.2518	.2373	−.0144
Jeep Grand Cherokee	25,849	.1475	.1376	−.010
Chevy S10 Blazer	22,651	.1106	.1071	−.0036
Toyota 4Runner	25,548	.0380	.0347	−.0033
Nissan Pathfinder	24,943	.0397	.0375	−.0022
Luxury cars*		.1610	.1565	−.0045
All vehicles		9.711	9.711	.000

NOTE.—Characteristics of the new car are described in the text.
* Cars priced above $30,000.

vehicles introduced are at the very upper end of the quality and price distributions of the SUVs offered in 1993; the Toyota SUV's price ($30,240) is $4,500 more than the most expensive SUV sold in 1993, and the Mercedes' price is $3,500 above that.

Table 12 summarizes results from introducing the Mercedes SUV. It did well, capturing about a third of the market share of the Explorer. The total number of vehicles sold hardly changed at all with the introduction; the demand for the Mercedes SUV comes largely at the expense of other SUVs and, to a far lesser extent, luxury cars. The Toyota SUV's introduction was somewhat less successful at our predicted price: its market share was only .05. To increase the Toyota SUV's market share to that of the Mercedes, we found that Toyota would have had to cut $1,000 off the price of its entrant. Our top predicted losers from the introduction of the Toyota SUV were the same as those for the introduction of the Mercedes SUV, but when the Toyota was introduced, the fall in the market share of luxury cars was much smaller. The Toyota Camry was the only nonluxury car that was in the top 15 of falls in sales, and it was in that list when either new SUV was introduced.

We cannot do a precise comparison of our out-of-sample predictions to the actual introduction of, say, the Mercedes M-Class SUV because there are many other confounded factors (the introduction of other new products and important macroeconomic shocks). However, we can note that the introduction of the Mercedes was generally considered to be very successful and was thought to put strong competitive pressure on other SUVs and on other luxury car makers (which is consistent with our prediction).

TABLE 13
DISCONTINUING THE OLDSMOBILE DIVISION

	Old Share	New Share	New − Old Share
All Oldsmobiles	.237	0	−.237
All GM	3.126	3.016	−.110
All cars	9.711	9.695	−.016
	Non-Olds Share Changes		
Chevy Lumina	.1354	.1548	.0194
Buick LeSabre	.1216	.1336	.0120
Pontiac GrandAm	.1322	.1441	.0119
Honda Accord	.2955	.3039	.0084
Ford Taurus	.2040	.2115	.0075
Saturn SL	.1465	.1539	.0074
Toyota Camry	.2343	.2415	.0072
Buick Century	.0614	.0683	.0069
Pontiac Grand Prix	.0517	.0584	.0067
Chevy Cavalier	.1700	.1767	.0067
Pontiac Bonneville	.0658	.0721	.0064

NOTE.—The original Oldsmobile models in the data (and their shares) are Ciera (.068), Cutlass Supreme (.059), Olds 88 (.050), Achieva (.033), Olds 98 (.019), and Bravada (.008).

B. Discontinuing the Oldsmobile Division

Table 13 provides the results from discontinuing the Oldsmobile division of General Motors. This is of interest because GM has in fact recently announced the phase-out of that division. In 1993 Oldsmobile had a market share of about 2.44 percent of the total number of vehicles purchased, whereas GM's total share of vehicles purchased was 32.2 percent. When we drop the Oldsmobile models from the choice set, the three vehicles that benefit the most are all family-sized GM cars (Chevy Lumina, Buick LeSabre, and Pontiac GrandAm). Still some of the Olds purchasers shift to high-selling family-sized cars produced by other companies, notably the Honda Accord, Ford Taurus, and Toyota Camry. Overall, 43 percent of Oldsmobile car purchasers substitute to a non-GM alternative, and GM's market share falls to 31.1 percent. Of course the profit change to GM depends on the costs saved by discontinuing Oldsmobile and on the markups of the GM cars that the Olds purchasers substitute to (numbers that GM presumably has detailed information on).[23]

VIII. Conclusion

In this paper, we explore the role of detailed consumer attribute data, together with second-choice data, in estimating a demand system for

[23] Since Oldsmobile is still in the process of shutting down, we cannot check our 1993-based estimates against what actually will happen. Of course there are also a number of other important changes in the market between 1993 and today.

passenger vehicles. We find that unobserved random coefficients are necessary to describe the relatively tight substitution patterns that are found in the data. The second-choice data are very helpful in obtaining precise estimates of the parameters that govern these substitution patterns. However, either some outside information or cross-sectional variation in choice sets must be used to pin down the absolute level of elasticities. As we have shown, these sources of data, when taken together, provide rich demand systems that imply realistic out-of-sample predictions.

Demand systems provide an important component of incentives for market responses to many (if not most) policy and environmental changes. We are hopeful that, given appropriate data, techniques that extend those provided here will enable researchers to analyze these changes in a useful way.

Appendix

Variances of Parameter Estimates

The variance-covariance of the parameters is determined by (i) the variance-covariance of the first-order conditions that define the estimator evaluated at the true value of the parameters and (ii) the expectation of the derivative, with respect to $\boldsymbol{\beta}$, of the first-order conditions that define the estimator evaluated at β_0 (see Hansen [1982] for the formula given these two matrices).

The variance in our moments when evaluated at θ_0 is generated by two sources of randomness: (1) sampling error in the CAMIP means (e.g., from the variance in $[n_j]^{-1}\sum_{i_j=1}^{n_j} z_{i_j}$) and (2) simulation error in our calculations of the model's predictions. Since the simulation and sampling errors are independent of each other and it is the difference between the sample mean and our model's predictions that enters our objective function (see eqq. [8] and [12]), the variance of the moment conditions can be expressed as the sum of the variances due to sampling and simulation errors. The variance due to sampling error can be consistently estimated by calculating the variance of the moment conditions at the estimate of the parameter values holding the simulation draws constant. The variance due to simulation error can be consistently estimated by simulating the sample moment at the estimate of $\boldsymbol{\beta}$ for many independent sets of ns simulation draws and calculating the variance across the calculated moment vectors.[24]

The derivative matrix can be consistently estimated by taking the derivative of the sample first-order condition evaluated at the estimate of $\boldsymbol{\beta}$, remembering that, since we use a two-step estimator, that derivative is the sum of two terms: one accounting for the direct effect of $\boldsymbol{\beta}$ on the moments given the estimate of $\boldsymbol{\delta}(\boldsymbol{\beta}, \cdot)$ and one accounting for the effect of $\boldsymbol{\beta}$ on $\boldsymbol{\delta}(\boldsymbol{\beta})$ (see, e.g., Pakes and Olley 1995).

[24] For each set of draws we have to solve the contraction mapping for the $\boldsymbol{\delta}^{N,ns}(\hat{\boldsymbol{\beta}})$ that corresponds to that set of draws and use that estimate of $\boldsymbol{\delta}^{N,ns}(\boldsymbol{\beta})$ in the calculation of the moments that go into (8) and (12). This is to account for the fact that the simulation affects both the prediction of the micro moments given an estimate of $\boldsymbol{\delta}(\beta_0)$ and the estimate $\delta_0(\beta_0)$, i.e. $\boldsymbol{\delta}^{N,ns}(\beta_0)$, itself.

References

Anderson, Simon P., André DePalma, and Jacques-François Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* Cambridge, Mass.: MIT Press.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *Rand J. Econ.* 25 (Summer): 242–62.

Berry, Steven T., James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (July): 841–90.

———. 1999. "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy." *A.E.R.* 89 (June): 189–211.

———. 2001. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." Working Paper no. 1337 (November). New Haven, Conn.: Yale Univ., Cowles Found.

Cosslett, Stephen R. 1981. "Maximum Likelihood Estimator for Choice-Based Samples." *Econometrica* 49 (September): 1289–1316.

Das, S., G. Steven Olley, and Ariel Pakes. 1995. "The Market for TVs." Working paper. New Haven, Conn.: Yale Univ.

Domencich, Thomas A., and Daniel McFadden. 1975. *Urban Travel Demand: A Behavioral Analysis.* Amsterdam: North-Holland.

Goldberg, Pinelopi Koujianou. 1995. "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry." *Econometrica* 63 (July): 891–951.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50 (July): 1029–54.

Hausman, Jerry A., and David A. Wise. 1978. "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences." *Econometrica* 46 (March): 403–26.

Heckman, James J., and James M. Snyder, Jr. 1997. "Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators." *Rand J. Econ.* 28 (special issue): S142–S189.

Hotelling, Harold. 1929. "Stability in Competition." *Econ. J.* 39 (March): 41–57.

Ichimura, Hidehiko, and T. Scott Thompson. 1998. "Maximum Likelihood Estimation of a Binary Choice Model with Random Coefficients of Unknown Distribution." *J. Econometrics* 86 (October): 269–95.

Imbens, Guido W., and Tony Lancaster. 1994. "Combining Micro and Macro Data in Microeconometric Models." *Rev. Econ. Studies* 61 (October): 655–80.

Lancaster, Kelvin. 1971. *Consumer Demand: A New Approach.* New York: Columbia Univ. Press.

Manski, Charles F., and Steven R. Lerman. 1977. "The Estimation of Choice Probabilities from Choice Based Samples." *Econometrica* 45 (November): 1977–88.

McCulloch, Robert E., Nicholas G. Polson, and Peter E. Rossi. 2000. "A Bayesian Analysis of the Multinomial Probit Model with Fully Identified Parameters." *J. Econometrics* 99 (November): 173–93.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by Paul Zarembka. New York: Academic Press.

———. 1981. "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data with Econometric Applications*, edited by Charles F. Manski and Daniel McFadden. Cambridge, Mass.: MIT Press.

McFadden, Daniel, et al. 1977. *Demand Model Estimation and Validation.* Berkeley, Calif.: Inst. Transportation Studies.

Nevo, Aviv. 2000. "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry." *Rand J. Econ.* 31 (Autumn): 395–421.
Pakes, Ariel. 1986. "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica* 54 (July): 755–84.
Pakes, Ariel, Steven T. Berry, and James H. Levinsohn. 1993. "Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of Environmental Change." *A.E.R. Papers and Proc.* 83 (May): 240–46.
Pakes, Ariel, and Steven Olley. 1995. "A Limit Theorem for a Smooth Class of Semiparametric Estimators." *J. Econometrics* 65 (January): 295–332.
Pakes, Ariel, and David Pollard. 1989. "Simulation and the Asymptotics of Optimization Estimators." *Econometrica* 57 (September): 1027–57.
Petrin, Amil. 2002. "Quantifying the Benefits of New Products: The Case of the Minivan." *J.P.E.* 110 (August): 705–29.