Lecture 3: NBERMetrics.

Ariel Pakes

July 17, 2012

Characteristics Based Models with Micro Data.

Micro data always matches individual characteristics to the products those individuals chose. This will provide information on

- the importance of interactions between observed individual characteristics and product characteristics.

In addition sometimes the data includes

- the characteristics of individuals that did not purchase one of the products being analyzed and this allows one to get a firmer grip on who values the outside alternative, and
- second choices (and in some contexts, like school choice, subsequent choices) by these same individuals, and as we shall see this is particularly useful as it gives us a handle on the importance of unobserved product characteristics.

Note. A related type of data is panel data which follows an individual's choices over time. The advantage of second choice data is it holds the household's conditions constant and gives an ordering of preferences given those conditions. When using panel data a household's condition can change over time and this can affect its preference ordering. However if one is willing to assume stable preferences over time the basic difference between panel data and second choice data is that one contains agents responses to a single change in choice set (the omission of the first choice), while the second typically records responses to a multitude of choice set changes.

Model. To begin we rewrite the model with explicit notation for the observed individual characteristics. If we let each individual have a different marginal utility from characteristic k we have

$$U_{ij} = \beta_0 + \Sigma_k \beta_{i,k} x_{jk} + \xi_j + \epsilon_{i,j}.$$

Now assume

$$\beta_{i,k} = \beta_k^o + \sum_r z_{i,r} \beta_{r,k}^o + \nu_{i,k},$$

where the $z_{i,r}$ are observed, the $\nu_{i,k}$ are not observed and can have different variances over characteristics (k). Notice that if we think of this just as a projection, then $\nu_{i,k}$ are that part of the variance of the coefficients that can not be accounted for by the observed $z_{i,r}$ and are uncorrelated with these observables by construction.

We can now rewrite the utility function in a more useable form as

$$U_{i,j} = \delta_j + \sum_{kr} x_{jk} z_{ir} \beta^o_{rk} + \sum_{kl} x_{jk} \nu_{il} \beta^u_{kl} + \epsilon_{ij}, \tag{1}$$

$$\delta_j = \sum_k x_{jk} \beta^o_k + \xi_j. \tag{2}$$

and the β^u_{kl} are the variances of the unobserved random coefficients on characteristics.

Interactions Between Individual and Product Characteristics. As noted in the earlier lectures interactions of consumer characteristics and product characteristics are needed for reasonable cross price elasticities. Now have two such interaction terms.

- Observed consumer characteristics (the z_i) and product characteristics
  (Term is $\Sigma_{kr} x_{jk} z_{ir} \beta^o_{rk}$.)

- Unobserved consumer characteristics (the ν_i) and product characteristics
  (Term is $\Sigma_{kl} x_{jk} \nu_{il} \beta^u_{kl}$.)

Types of Micro Data.

- random sample of consumers
- choice-based sample; samples of purchasers of particular products (sometimes purchasers of one of the products) and find out their characteristics.

- Typical kind of data from marketing firms as this is a more efficient way of getting the information they need. Sometimes they let the actual micro data out, but more often the only variables that are available are the average characteristics of consumers who purchased particular goods. Also often the only data available is older data, data that they are not currently marketing.
- To use this data we often need a correction for "choice-based" sampling.

Adding Information on The Correlation Between Consumer and Product Characteristics to Product Level Data.

Often all we know in addition to the market level data is either the average characteristics of consumers who chose a particular product, or the covariance of product and consumer characteristics. The easiest way to use this data is to add moments to the market level data estimation algorithm. That is we use the model to predict this correlation conditional on a given parameter vector, and then add a moment which is the difference between the model's prediction for the correlation conditional on θ and the correlation in the data.

Since the precision of the moment will depend on n the number of individuals underlying the consumer samples' averages, ns the number of simulation draws we use (see below), and N the size of the market underlying the market level data, we index the added moment by all three of these variables. We let $x_{k,j}$ be the value of the k^{th} characteristic of the j^{th} car and $y_i^1 = j$ symbolize the event that consumer i's first choice was product j. Then our moment is written as

$$G_{n,ns,N}(\theta) = \Sigma_j \frac{n_j}{n} x_{kj} \left[n_j^{-1} \Sigma_{i_j=1}^{n_j} z_{i_j} - E[z|y^1 = j, \theta] \right]$$

where $\theta = (\beta, \delta)$. The only component of this expression that one needs the model's implications for is $E[z|y^1 = j, \theta]$. Its actual value is hard to compute, but when we transform it using Bayes rule it is easy to approximate. In particular

$$E[z|y^1 = j, \theta] = \int_z z\mathcal{P}(dz|y^1 = j, \theta)$$

$$= \frac{\int_z zPr(y^1 = j|z, \theta)\mathcal{P}(dz)}{Pr(y^1 = j, \theta)}$$

$\mathcal{P}(dz)$ denotes the density of z in the market of interest. We will assume this is available from the CPS or some other survey. Now

$$\frac{\int_z zPr(y^1 = j|z, \theta)\mathcal{P}(dz)}{Pr(y^1 = j, \theta)} \approx$$

$$\frac{\int_z zPr(y^1 = j|z, \theta)\mathcal{P}(dz)}{s_J^N} \approx$$

$$\frac{(ns)^{-1} \Sigma_r z_r p\left(y^1 = j|z_r, \nu_r, \beta, \delta^{N,ns}(\beta)\right)}{s_j^N},$$

where the first approximation is because $\{s_j^N\}$ the vector of observed market shares will equal the model's prediction by choice of $\{\delta\}$ for each θ. Since (z_r, ν_r) for $r = 1, \ldots, ns$ index ns random draws on a couple whose first component, z_r, is taken from the CPS and whose second component, ν_r, is taken from the assumed distribution of ν given z, the second approximation follows from the properties of simulation estimators.

Petrin on the Minivan; JPE 2001

History: Don DeLaRussa was a head design engineer with Ford who had the idea of building a car with lots of passenger space but drove like a family car (had front wheel drive, proper suspension, ect.). The competitors to such a car at the time were station wagons, the Volkswagon Microbus, and some other full size vans that were difficult to drive. Ford decided not to do it, partly because of a worry that it would cannibalize its station wagons (it had a disproportionate share of that market).

Iacocca moves from Ford to Chrysler (1979), and brings DeLaRussa with him. At the time Chrysler looked like it might go under, and changes had to be made. Moreover, they had a small part of the station wagon market. They developed the Minivan; first marketed with the Dodge Caravans and Plymouth Voyagers in 1983. There are varying estimates of development cost, but almost all of them are below one billion 1983 dollars (about 3/4 of a billion seem to be the best guess).

The Minivan was a fairly immediate success. GM and Ford misread the success as demand for a business vehicle and hence introduced the Astro (GM) in 1985 and the Ford Aerostar in 1986; but these were rear wheel drive vehicles that handled more like a truck. It took until 1990 for GM and Ford to introduce a real competitor (three or four years from the time they realized what they wanted; which is typical lead time in this industry). By 1988 Chrysler had already sold over 1.3 million Minivans for over ten billion in sales; and seeing they had few competitors in this side of the market, markups were high. Chrysler went from being the "failing" car company, to the most profitable. Moreover Chrysler kept improving its Minivan, and maintained the lead in that market (which is over 10% of new vehicle sales) well into the mid 1990's.

The Goal and Its Importance. The goal is to quantify the impact of the introduction of the Minivan on; (i) Chrysler's profits, (ii) Consumer Welfare, and (iii) on the profits of competing firms. It is the comparison between (i) and (ii) that leads to the difference between the social and private returns to research that it is the underlying reason for various subsidies and tax breaks for R&D, and much of the growth literature in economics. Of course this was one successful innovation, and one can not generalize from it to the difference between social and private returns without considering also unsuccessful innovations and the appropriate way of building a sample of innovations.

The Exercise Estimate a demand and cost side from the observed data, and then recompute equilibrium prices and quantities from a choice set that does not include the minivan.

Conceptual problem's that arise;

- Were the minivan not have been developed other car prices would have re-adjusted due to a change in competition. If we can compute a new Nash pricing equilibrium and it is unique, we can correct this problem.
- Had the minivan not been developed, different new vehicles would have been introduced over the time period studied. This we can not correct without a dynamic model of characteristic choice. Dynamic models in I.O. are in a more primitive state of development, but we are much closer to being able to use them for questions like this (see for e.g. Doraszelski and Pakes' review, 2007, *Handbook of I.O.*)

Of course one has to start the analysis of these issues somewhere, and this is one component.

To the standard market level data, Petrin adds first-choice moments obtained from the consumer expenditure survey (CEX). The CEX is the data set that underlies the weights used for different commodity groups in the CPI. It is too small to use in an econometric model to directly analyze consumer choice (there are many vehicle models which are not chosen by anyone in the survey). However Petrin uses it to add micro-data based moment restrictions

- One set corresponding to purchase probabilities interacted with individual characteristics and
- one set of car characteristics interacted with individual characteristics.

Other than this the methods are those described in our previous lectures, so I go straight to the conclusions.

Aviv went over the various estimates and welfare calculations. I will quickly comment on two of the economics implications of the results. I want to just quickly review its implications for prices and the comparison between the private and social returns to research.

Prices. Table 7 gives some of the differences in prices between the two equilibria. Notice that Chrysler actually increases the prices of its family cars, since now if households decide to leave those cars due to the price increase they are likely to go to another Chrysler product, the Minivan. However almost all the prices of Ford and GM family cars fall, as they now have a new potent competitor.

TABLE 7
EQUILIBRIUM PRICES WITH AND WITHOUT THE MINIVAN, 1984: 1982–84 CPI-ADJUSTED DOLLARS

	Price: With Minivan	Price: Without Minivan	ΔPrice	% ΔPrice
	A. Largest Price Decreases on Entry			
GM Oldsmobile Toronado (large sedan)	15,502	15,643	−141	.90
GM Buick Riviera (large sedan)	15,379	15,519	−139	.89
GM Buick Electra (large sedan)	12,843	12,978	−135	1.04
GM Chevrolet Celebrity (station wagon)	8,304	8,431	−127	1.51
Ford Cadillac Eldorado (large sedan)	19,578	19,704	−126	.64
Ford Cadillac Seville (large sedan)	21,625	21,749	−125	.57
GM Pontiac 6000 (station wagon)	9,273	9,397	−123	1.31
GM Oldsmobile Ciera (station wagon)	9,591	9,714	−123	1.27
GM Buick Century (station wagon)	8,935	9,056	−121	1.34
GM Oldsmobile Firenza (station wagon)	7,595	7,699	−104	1.35
	B. Largest Price Increases on Entry			
Chrysler LeBaron (station wagon)	9,869	9,572	297	3.10
Volkswagen Quattro (station wagon)	13,263	13,079	184	1.41
Chrysler (Dodge) Aries K (station wagon)	7,829	7,659	170	2.22
AMC Eagle (station wagon)	10,178	10,069	109	1.08

NOTE.—Equilibrium prices without minivans are estimated using the model with microdata and Bertrand-Nash first-order conditions. Bertrand-Nash pricing with random coefficients does not a priori determine signs of firm-specific price changes.

Welfare, Profits and The Returns to Research. Table 13 gives us total change in Welfare from the introduction of the Minivan. Note that the compensating variations here, that is the income change that would make each consumer just indifferent between the choice set without the minivan and the choice set with the minivan, consist of;

- the gain to consumers who purchased a minivan, which, assuming preferences do not change over time, includes
  - the logit estimate of the infra-marginal returns to consumers who purchased a Minivan at its high price (this is just an out of sample extrapolation whose value depends on the assumed functional forms), and
  - the gain to those who did not purchase the good at some price but did at a lower price (with enough data one could get bounds on this without parametric assumptions),
- the change in utility to consumers who did not purchase a minivan but saw the prices of other vehicles change.

There were profit losses due to the minivan for Ford and GM in each of these years. See Table 11. So all the profits (and much more) are Chryslers. Still over three billion in welfare gains went directly to consumers and was not captured by Chrysler. If it was really a 1 billion investment, the excess social rate of return was on the order of 50% per year. Profit figures not shown here imply that the return to Chrysler over those years was slightly under $35-45\%$ per year.

TABLE 13
CHANGE IN U.S. WELFARE FROM THE MINIVAN INNOVATION, 1984–88 ($ Millions)

Year	Compensating Variation	Change in Producer Profits	Welfare Change
1984	367.29	−36.68	330.61
1985	625.04	−25.07	599.97
1986	439.93	27.30	467.23
1987	596.59	29.75	626.34
1988	775.70	110.24	885.94
Total	2,804.55	105.54	2,910.09

NOTE.—Computations were done using 1982–84 CPI-adjusted dollars.

TABLE 11
CHANGE IN INDUSTRY AND BIG THREE TOTAL VARIABLE PROFITS WITH THE ADVENT OF MINIVANS

YEAR	INDUSTRY	CHRYSLER		FORD		GM	
1984	−.21%	$202.5	14.38%	−$31.8	−1.16%	−$155.8	−1.50%
1985	−.13%	$259.1	13.99%	−$37.4	−1.29%	−$171.0	−1.63%
1986	.14%	$201.1	12.42%	$54.7	1.84%	−$119.9	−1.09%
1987	.17%	$346.1	23.27%	−$22.8	−.66%	−$174.5	−2.14%
1988	.65%	$504.1	32.50%	−$24.7	−.70%	−$235.4	−2.90%

NOTE.—Dollar figures are given in millions. The numbers are computed using the model to estimate profits both with minivans in the market and with minivans removed from the market (see Sec. V).

It is these kind of numbers, particularly the excess returns numbers, that get the profession excited about new goods, R&D, and growth; but remember this was a highly successful innovation. Maybe one or two innovations like this have occurred in the car industry over the last twenty years, and they do a lot of R&D. So this is just one component of what should be a much more complex calculation.

Combining Consumer Level and Aggregate Data.

We now consider individual consumer choice data directly in conjunction with aggregate data. The presentation will follow MicroBLP (JPE, 2005), as that paper has both first and second choice data, and what each source of information can be used for. The micro data The CAMIP data; which is proprietary *General Motors Corporation's* data used for their internal marketing and product quality programs.

In summary the data that we use has

- vehicle characteristics, prices, and sales (similar to product level data already in use except of higher quality)
- household characteristics by vehicle purchased (age, income, family size, ... broken down by vehicle purchased)
- second choice vehicles (generated as the reply to the question: "If you did not purchase this vehicle, what vehicle would you purchase?")

Note however that unlike most studies on aggregate data, we only have information on one market in one period.

General Points.

Recall from equation (1) that

$$U_{i,j} = \delta_j + \sum_{kr} x_{jk} z_{ir} \beta^o_{rk} + \sum_{lk} x_{jk} \nu_{il} \beta^u_{lk} + \epsilon_{ij}.$$

Here are some general points to keep in mind.

- The benefit of having the "first choice" micro data is that it matches observed individual attributes to the products those individuals chose. Consequently, the contribution of that data will be determined by the importance of the observed attribute data in determining individual demands (by the β^o_{rk}, and the variance of the z_i). Using our notation if $\beta^o_{rk} \simeq 0$ the aggregate purchase proportions (i.e. aggregated data on sales and characteristics) are sufficient statistics for the micro first choice data (i.e. we revert to BLP's problem).

- The unobserved individual attributes (the ν_i) differentiates this model from the "standard" micro data based logit model (models with no random coefficients) and focuses attention on two issues.

  - A substantive issue, i.e. what is the "quality" of the *observed* individual attribute data (are most of the attributes that cause different preference intensities for different product characteristics observed?)
  - A computational issue. If the ν_i are important the individual choice probabilities have to be simulated, and this complicates the estimation algorithm. Leads to desire to "Test" the null $\beta^u_{lk} = 0$.

- When there is only data on one market in one period there is little intuition leading us to expect that the first choice data will help us identify the impact of unobserved individual characteristics on preferences (since neither the choice set nor the distribution of consumer characteristics, changes over subsets of the data). However the unobserved characteristics *do affect* the relationship between first and second choices (regardless of the importance of observed attributes). Thus we expect to be able to engage in a far more detailed investigation of the importance of unobserved differences in preference intensities once we add the second choice data to the combination of data sources we are using.

Unobserved Product Characteristics & Micro Data. The micro data enables us to estimate a separate constant term for each choice, our

$$\delta_j = \Sigma_k x_{jk} \beta_k^o + \xi_j,$$

and the vector

$$\beta \equiv (\{\beta_{r,k}^o, \beta_k^u\}_{r,k})$$

I.e. β can be estimated without assuming anything about the distribution of ξ_j.

However, since

$$\delta_j = \Sigma_k x_{jk} \beta_k^o + \xi_j, \tag{3}$$

we need β to compute elasticities w.r.t. the product attributes (including price). Different assumptions on the joint distribution of the observable characteristics and the $\{\xi_j\}$ would allow us to estimate the β_k^o, but the assumptions that would fully identify the model when we add the micro data are no different than the assumptions that would identify it when only macro data are available.

Note. There are issues we can address without estimating the $\{\beta_k^o\}$; e.g. the analysis compensating variation as we can hold the $\{\delta_j\}$ and still analyze that. However any issue that we want to address that requires a response to a change in value of any characteristic (price, or any other) will require estimates of β_k^o, and the essence of the problem of getting these coefficients is similar to the problem we faced when we used only aggregate data to estimate coefficients.

Choices in Estimation and Computation.

There are at least two choices of what to estimate. We could either

- estimate (β, δ) pointwise, or
- make assumptions on the distribution of ξ (eg. $E[\xi|x] = 0$), and estimate (β, β^o) instead of (β, δ).

The tradeoff here is clear enough. By estimating (β, β^o) instead of (β, δ) we gain efficiency if the assumption on the properties of ξ are correct, but loose consistency if it is wrong. Typically the choice depends on the nature of the data available. If there is a lot of consumer level data you might want to estimate (β, δ), and then perhaps use those estimates to examine the robustness of alternative assumptions on ξ. This is what MicroBLP does.

Computational Choices: Search Routines. Again there are at least two ways to proceed. One could

- search over (β, δ).
- or use a two step estimator.

The two step estimator would use the contraction mapping in BLP and product level data to solve for δ as a function of (β, s^N, P^{ns}), i.e. for the $\delta^{N,ns}(\beta)$ that solves

$$s(\beta, \delta, P^{ns}) = s^N$$

then search over the other parameters (β) to match the model's predictions to the micro data.

Again the choice will typically depend on the nature of the problem and the data. In MicroBLP there were over 200 products (and each one has a δ) in addition to the β. That many dimensional search is pretty hard to solve, though the new search routines mentioned in prior lectures definitely make this easier, it is still probably a little heroic to institute a search over all parameters (though this would be the efficient way to go). So MicroBLP chooses to use the second method, and a first step which estimates $\delta^{N,ns}(\beta)$.

Computational Choices: Handling The Micro Data. Assume now that we are doing the first search (joint over (β, δ)) as we will see that what I am about to say applies, *a fortiori* to the second search procedure. Probably one's first thought for estimating these parameters is to go to mle. However, that often leads to very poor results.

MLE is only efficient when one can calculate the likelihood probabilities exactly. If you allow for random coefficients you will have to approximate the likelihood. If you use simulation you do not construct $p(z_i|\beta,\delta)$ but $\hat{p}_{ns}(z_i|\beta,\delta)$. These probabilities enter the likelihood inside a logarithm so using an expansion to evaluate the approximation and letting

$$u_{ns}(z_i|\beta,\delta) = \hat{p}_{ns}(z_i|\beta,\delta) - p(z_i|\beta,\delta)$$

we have

$$log[\hat{p}_{ns}(z_i|\beta,\delta)] - log[p(z_i|\beta,\delta)] \approx$$
$$\frac{u_{ns}(z_i|\beta,\delta)}{p(z_i|\beta,\delta)} + \frac{u_{ns}(z_i|\beta,\delta)^2}{2p(z_i|\beta,\delta)^2} + o(p(z_i|\beta,\delta)^3).$$

Provided we use an unbiased simulator the first term has expectation zero and since we will be averaging over lots of mean zero errors, it will not be a problem. However the second term's expectation is

$$\frac{\sigma^2_{ns}(z_i|\beta,\delta)}{2p(z_i|\beta,\delta)^2} \approx \frac{1}{2 \times ns \times p}.$$

With enough ns this will also go to zero, but think of how large a problem this is. In most markets most individuals do not chose a product in every period. In the auto market about one in every 10 households chose a car in each year. Often there are many products to split these 10 households probabilities into. In MicroBLP there were 200 products so it would not be unusual to have probabilities on the order of .0001. That is, to get the variance just down the order of magnitude of the probability, we would need over 10,000 random draws per individual, and there were about 30,000 individuals. Moreover, since the variance is a function of (β, δ), if you do not get rid of it, it will bias the parameter estimates. This problem becomes even worse when we use the second search procedure for then one has to account also for the error in the estimated δ's.

This is a rather extreme case, but quite frequently it is not sensible to use maximum likelihood for these problems. This is because the logarithmic function exacerbates the errors at small p. There are a number of alternatives (including more precise ways of obtaining estimates of the integral defining the probabilities), but for large problems it makes sense to move to an estimator which is linear in the simulation errors. Then the estimator will average out errors over individuals in the consumer sample.

MicroBLP's choice of moments. They chose moments that have an obvious interpretation in terms of fit. CAMIP contains random samples of size n_j from the set of households who chose car j (for $j = 1, \ldots, J$). Each sampled households reports its characteristics and the vehicle it would have chosen had its first choice not been available. So we fit the following covariances.

- The covariances between household characteristics and first choice car characteristics (eg. the covariance between family size and car size). Use one such covariance for every interaction between observed individual characteristic (the z_i) and car characteristic (the x_{kj}) appearing in the utility function (in U_{ij}) in equation (3) and interactions between z_i and a constant term. The latter determines the average characteristics of the households in the CAMIP sample (households that buy cars), and hence U_{i0}. These moments are calculated as in the derivation given earlier in this lecture.

- The covariances of the characteristics of the first and second choice cars (eg. the covariances between the size of the first and second choice cars of the same households). One such covariance for each car characteristic which is a determinant of U_{ij}. (Analogously our utility specification will allow for one unobserved taste shifter for each observed product characteristic). Intuitively these moments allow us to determine the impact of *unobserved* consumer attributes on preferences for car characteristics and for the outside good (the β^u).

Summary. The nested method of moments algorithm used in MicroBLP was as follows.

- First step. As in BLP we first find, for any given β, the value of δ which makes the aggregate shares derived from the model fit the product level demands

$$s(\beta, \delta, P^{ns}) = s^N \ \rightarrow \ \delta^{N,ns}(\beta).$$

- Second step. Substitute $\delta^{N,ns}(\beta)$ for δ, into the model, and compute the model's predictions for the micro data moments as a function only of β. Select β that minimizes a distance between the moments predicted by the model and the moments in the data.

Data and Tables.

CAMIP is obtained as a random sample of households who registered particular new cars in 1993, and is rather large (after excluding observations with missing attribute data, we still have 37,500 observations, 34,500 of which also have second choice data.) The number of observations per vehicle was determined by GM (and is known to us).

CAMIP vs the CPS. Table 1 compares the distribution of household characteristics in the CAMIP sample to those in the CPS. Recall that the CAMIP sample is close to a random sample from those who purchased a car. One way to look at this comparison is as the magnitude of the issue induced by choice based sampling. Clearly the CAMIP sample has more income per household than the CPS sample, and is significantly less urban and more rural (apparently the rural population purchases a disproportionate number of vehicles, a fact which might help explain the share of trucks in total vehicle sales; see below).

Some Details.

The Choice Set. We were given the CAMIP sampling cells (basically make, model, and body style). 200 choices, a quarter of them light trucks (Minivans, SUV's,FSVans, light trucks) which account for about 40% of sales in 1993. We have price and options for every car; use modal characteristics within a cell, and average price at which the modal characteristics sold for models.

Characteristics of Data. Total vehicle sales about 196 billion dollars, and total light truck sales about 81 billion dollars. Table 4 attempts to figure out which individual characteristics might be interacted with which vehicle characteristics (even this data set will not support all possible interactions). All your intuitive notions of what should be correlated with what seem to be supported by the data. Table 5 provides some feeling for the second choice data. It is generated as the response to the

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

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 6
ESTIMATES OF INTERACTION TERMS, β^o

VEHICLE CHARACTERISTIC AND HOUSEHOLD ATTRIBUTE	FULL MODEL (1)	LOGIT: First (2)	LOGIT: 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)

question:*Suppose the vehicle you bought were not for sale. What other vehicle would you have purchased instead?* There are quite a number of 2^{nd} choices for each first choice car; though the second choices are more concentrated for light trucks (indicating that this is where you expect to see higher markups and, characteristically, this is where you have seen entry since 1993). All first and second choice characteristics correlations are positive, and highly significant (though price, minivan, sport utility, pickup, are larger than the others)[1]

Estimates Presented. We provide several sets of estimates for comparison; i) from our model, ii) from a model that does not allow for any interactions between observed product characteristics and unobserved consumer characteristics (the *logit model*);once using only first choice data (labeled Logit 1^{st}; use MLE), and once using both first and second choice data (labeled Logit 1^{st}&2^{nd} choice; also MLE.); iii) a model which does not allow any interactions between observed product and observed consumer characteristics, a sigmas only model (were we to use first choice data only, this would be identical to using only aggregate data, as in BLP).

Specification Used. In the specification discussed here our interaction terms have for all variables except price;

$$\beta_{ik} = \Sigma_l z_{il} \beta^o_{lk} + \nu_{ik}.$$

We were more detailed with the price interaction. We assumed the coefficient of price was $exp[-W_i]$ where W_i or "wealth" is determined as

$$\text{spline in income} + \beta \ \text{ family size} + \nu_{ip}$$

where the $\{\nu_{ik}\}$ are i.i.d. both over characteristics for a given individual and over individuals.

The Estimates.

- *Observed interactions: Table 6.* Pretty much all variables are of the expected sign; and the logits are not terribly different, in either sign or significance, then the results from the full model. They do have some anomalies; notably in their coefficients for the outside good and, for the first choice logit, the negative rural pickup interaction.

[1] It is interesting to note that with second choice data you can have a direct test of the IIA property by looking at second choices of individuals with different initial choices. It clearly fails.

TABLE 6
ESTIMATES OF INTERACTION TERMS, β^o

VEHICLE CHARACTERISTIC AND HOUSEHOLD ATTRIBUTE	FULL MODEL (1)	LOGIT First (2)	LOGIT 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)

• *The* β^u*; Table 7.* Out of 23 interactions at least 20 are significant, several with t-values over twenty. This is an overwhelming rejection of: $H_0 : \beta^u = 0$, or that there are no unobserved household attributes which interact with vehicle attributes. These results are even more striking for the "sigma only" model.

• *Fit.* The logits provide an adequate fit for the correlations between observed household characteristics and observed vehicle characteristics, but do very poorly in matching the characteristics of the first and second choice car. The "sigmas only" model 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 observed car characteristics. The "Full Model" does about the same as the best of the alternatives in *both* dimensions.

Details. Income has two independent effects. It increases the value of the outside good, so families with high income have to obtain a higher utility from purchasing before they purchase than those with low income. The interaction between income and price tells us that they do, especially for high priced cars. As expected, the exact opposite holds for family size conditional on income (larger families behave as if they are poorer than smaller families with the same income). Families with children have a strong preference for minivans, though somewhat surprisingly children do not have a noticeable impact on the size of non minivan cars. Adults do however. Age is one of the more important correlates of preferences, interacting significantly with several car characteristics. Rural families have a distinct preference for pickups and sport utility vehicles. There seems to be a stronger "brand" preference for the U.S. automakers than for there Japanese counterparts,....

TABLE 7
ESTIMATES OF INTERACTION TERMS, β^u

Parameter Name	Full Model (1)	$\beta^u \equiv 0$ (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.

Estimating the β^o

Recall that

$$\delta_j = \Sigma_k x_{jk}\beta_k + p_j\beta_p^o + \xi_j.$$

- and we expect p_j to be correlated with ξ_j, so we need instruments to obtain consistent estimates of the parameters of this equation
- Unlike BLP, who had twenty cross-sections with which to estimate this equation, we only have the data for 1993. This suggests a precision problem similar to BLP's; but this time only for a subset of the parameters of interest (the β^0).
- We are particularly interested in β_p^o, since it plus the other parameters already estimated determines all own and cross price elasticities.

Possibilities that enable more precise estimates of β_p^o

- add an assumption on marginal cost and a pricing equation as in BLP or
- we can find prior information on β_p^o. We try two sources. The value which generates an aggregate elasticity of one (which was GM's estimate of the aggregate elasticity) and zero, as that is the implicit assumption in models that ignore the endogeneity problem when using micro data.

We conclude from tables 8 that levels of elasticities differ markedly with the estimate of β_p^o used but the pattern of them does not

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

(the different estimates just shift them all up and down together). Note that as might have been expected from looking at the second choice data, elasticities are particularly low for light trucks (as there were few competitors). This means markups are higher here, and we might expect the addition of new vehicles in these "niches"; which is exactly what happened.

—Large Interpreting the Results. Note that the parameters per se are not very interesting. The major interest is in their implications. The major implications we investigate are their implications for price substitutes, and for product substitutes (Tables 9 and 10). These two differ because when price goes up only a selected sample of people leave that car, those purchasers who were particularly price sensitive, while when the product is pulled out all people who purchased that car must move.

These tables show you just how important it is to have allowed sufficient room for unobservable interactions.

Prediction Exercises.

Though getting an accurate indication of price effects on demand is probably the single most important use of demand systems, demand systems are used in other ways and this section illustrates two such uses:

1. Evaluating the potential demand for new models; in particular we introduce "high-end" sport utility vehicles (SUV).

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.

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

Vehicle	Full Model	Logit: First	Logit: 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

2. 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 sport utility vehicles were introduced in the late 1990's (an apparent response to the high markups being earned on those vehicles in the period of our data; see Table 7), and GM announced its intention to close down its Oldsmobile division in 2000.

Two caveats

- First of all, the data used in our investigations is 1993 data. The market has changed since 1993 and those changes might well affect the results.
- Second, in the exercises done here we do not allow other actors in the market to respond to the change we are investigating. I.e. when we shut down the Oldsmobile division, we do not allow for either a realignment of the prices of other products in response to the shutdown, or for the introduction of the new models that might follow such a shut down. Similarly when we introduce a new model 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. To do price responses, we would need a cost system and a pricing equation. To do more we would need a dynamic model.

New Models.

The two new models we introduce 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 sport utility vehicle at that time.

Recall that ξ captures the effect of all the detailed characteristics that are omitted from our specification; we think of it as "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[2]. Both vehicles introduced are at the very upper end of the quality and price distributions of the SUV's offered in 1993; the Toyota SUV's price ($\$30,240$) is $\$4,500$ more than that most expensive SUV sold in 1993 and the Mercedes' price is $\$3,500$ above that.

The table summarizes results from introducing the Mercedes

[2]The mean Mercedes quality and price 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 doing prediction way out of the range of the data which we used in our estimation (and probably also out of the range of the SUV eventually marketed by 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 sports utility vehicles, and to a far lesser extent, from four 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 a thousand dollars off the price of its entrant. Interestingly when Toyota did enter this market it entered with a high end SUV (the Lexus) which was more similar to the Mercedes as well as with the Toyota SUV.

Our top predicted losers from the introduction of the Toyota SUV were the same as those for the introduction of the Mercedes SUV. However, when the Toyota was introduced the fall in the market share of luxury cars was not noticeable. The only non-luxury car which was hurt by either new SUV was the Toyota Camry, and it was among the top 15 losers from both introductions.

Discontinuing the Oldsmobile Division.

The table provides the results from discontinuing the Oldsmobile division of *GM*. In 1993 Oldsmobile had a market share of about 2.44% of the total number of vehicles purchased, while *GM*'s total share of vehicles purchased was 32.2% . When we eliminate Oldsmobile models in our data set (see the bottom panel of Table 2), the three vehicles which benefit the most are

Introducing a Mercedes SUV.*

Model	Price	Old Share	New Share	New - Old Share
NEWCAR	33.659	0.0000	0.0762	0.0762
Biggest Declines in Sales.				
FORDEXPL	24.2740	0.2518	0.2373	-0.0144
JPGRNDCH	25.8490	0.1475	0.1376	-0.010
CHS10BLZ	22.6510	0.1106	0.1071	-0.0036
TO4RUNNR	25.5480	0.0380	0.0347	-0.0033
NISSPATH	24.943	0.0397	0.0375	-0.0022
Luxury cars**	top 4	.1610	.1565	-.0045
All Vehicles	n.r.	9.711	9.711	.000

* See the text for the characteristics of the new car.
** Cars priced above $30,000.

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 the Toyota Camry. Overall, 43% of Oldsmobile car purchaser substitute to outside of GM, and GM's market share falls to 31.1%. 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 which GM presumably has extensive information on).

Discontinue 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.			
CHELUM	0.1354	0.1548	0.0194
BUILES	0.1216	0.1336	0.0120
PONGAM	0.1322	0.1441	0.0119
HONACC	0.2955	0.3039	0.0084
FORTAU	0.2040	0.2115	0.0075
SATSL	0.1465	0.1539	.0074
TOYCAM	0.2343	0.2415	0.0072
BUICEN	0.0614	0.0683	0.0069
PONGRA	0.0517	0.0584	0.0067
CHECAV	0.1700	0.1767	0.0067
PONBON	0.0658	0.0721	0.0064
Olds Share Changes			
OLDCIE	0.0676	0.00	-0.0676
OLDSUP	0.0594	0.00	-0.0594
OLDn88	0.0498	0.00	-0.0498
OLDACH	0.0333	0.00	-0.0333
OLDn98	0.0187	0.00	-0.0187
OLDSBRAV	0.0083	0.00	-0.0083