Static Models of Demand for Differentiated Products

We previously motivated the importance of allowing for a flexible model of demand in order to infer conduct. We will now cover: the difficulties in modeling demand in differentiated-products markets, and solutions offered in the literature.

The Problem

The most straight-forward approach

$$q = D(p; r)$$

where: q – a J-dimensional vector of quantities demanded;
p – a J-dimensional vector of prices;
r – a vector of exogenous variables;

The main concern of previous work was to specify $D(\cdot)$ in a way that was both flexible and consistent with economic theory.
Linear Expenditure model (Stone, 1954); Rotterdam model (Theil, 1965; and Barten 1966); Translog model (Christensen, Jorgenson, and Lau, 1975); Almost Ideal Demand System (Deaton and Muellbauer, 1980);

For differentiated products the problems with these:

1) Dimensionality: due to the large number of products the number of parameters will be too large to estimate.
   Example: a linear demand system, $D(p) = Ap$, where A is $J \times J$ matrix of constants, implies J^2 parameters. This problem is augmented if we attempt to use a flexible functional form.

2) Multicollinearity/IV – In most differentiated products industries prices of the various goods will be highly collinear. This problem is augmented since we require an IV for each price. It is usually very hard to find IV that are both exogenous and will not generate moment conditions that are not nearly collinear.
3) Heterogeneity – for some applications we would like to explicitly model and estimate the distribution of heterogeneity.

Symmetric Representative Consumer Models

These models solve the dimensionality problem imposing symmetry assumptions that restrict the patterns of substitution.

Example 1: Constant elasticity of substitution (CES) utility:

$$U(q_1,\ldots,q_J) = \left(\sum_{i=1}^{J} q_i^{\rho} \right)^{1/\rho},$$

where ρ is a constant parameter that measures substitution across products. The demand of the representative consumer obtained from this utility function is

$$q_k = \frac{p_k^{-1/(1-\rho)}}{\sum_{i=1}^{J} p_i^{-\rho/(1-\rho)}} I, \quad k = 1,\ldots,J,$$

where I is the income of the representative consumer.

The dimensionality problem is solved by imposing symmetry between the different products; thus, estimation involves a single parameter.

This model implies

$$\frac{\partial q_i}{\partial p_j}\frac{p_j}{q_i} = \frac{\partial q_k}{\partial p_j}\frac{p_j}{q_k}, \text{ for all } i,\ k,\ j.$$

The cross-price elasticities are restricted to be equal, regardless of how "close" the products are in some attribute space.

This model is popular in theory and is sometimes used in trade and in macro, but it is not appropriate for our purpose.

Example 2: An alternative to the CES utility function is

$$U(q_1, ..., q_J) = \sum_{j=1}^{J} \delta_j q_j - \sum_{j=1}^{J} q_j \ln q_j \ ,$$

which yields the Logit demand.

The utility function for the Logit representative consumer has two terms. The first suggests that the representative consumer will consume only the product with the highest δ_j. The second term is an entropy term and expresses a variety-seeking behavior. Through this second term we get consumption of more than one product, but its functional form illuminates the problem all products enter this entropy term in a symmetric way.

Estimation of this model involves J parameters and allows for somewhat richer substitution patterns.

As we discuss below the substitution patterns in the Logit model are solely a function of market shares (which here are equivalent to the quantities consumed by the aggregate consumer), and are not related to the characteristics of the products.

Multi-Stage Budgeting

The basic idea: solve the dimensionality problem by dividing the products into smaller groups and allow for a flexible functional form within each group.

Multi-stage budgeting – the consumer maximization problem is split into several stages:
– at the highest stage expenditure is allocated to broad groups;
– at lower stages group expenditure is allocated to sub-groups;
– At each stage the allocation decision is a function of only that group total expenditure and prices of commodities in that group (or price indexes for the sub-groupings).

There are various conditions that will guarantee that the solution to this multi-stage process will equal the solution to the "full" problem.

Weak separability of preferences:

$$U(q_1,q_2,...,q_J) = f[v_1(q_1,q_2),v_2(q_3,q_4),...,v_G(q_{J'},...,q_J)],$$

where $f(\cdot)$ is some increasing function and $v_1,...,v_G$ are the sub-utility functions associated with separate groups.

Weak separability is necessary and sufficient for the last stage of the multi-stage budgeting; if a subset of products appears only in a separable sub-utility function, then the quantities demanded of these products can always be written as only a function of group expenditures and prices of other products within the group.

In higher stages usually rely (1) indirect utility functions for each segment are of the Generalized Gorman Polar Form, and (2) the overall utility is separable additive in the sub-utilities.

Originally developed for the estimation of broad categories of products.

Hausman, Leonard, and Zona (1994) and Hausman (1996) use the idea of multi-stage budgeting to construct a multi-level demand system for differentiated products.

Example: Beer (HLZ) or Cereal (Hausman).
Upper level: category demand
Middle level: segment demand
Lower level: brand demand

Each level allows for a flexible functional form;

Assume the data is for $j = 1, \ldots , J$ products in $t = 1, \ldots, T$ markets:

Lowest level : the Almost Ideal Demand System

$$s_{jt} = \alpha_j + \beta_j \log(y_{gt}/\pi_{gt}) + \sum_{k=1}^{J} \gamma_{jk} \log p_{kt} + \epsilon_{jt},$$

where: s_{jt} is the dollar sales share of total segment expenditure,
y_{gt} is overall per capita segment expenditure,
π_{gt} is the price index;
p_{kt} is the price of the kth brand in market t.

The price index, P_{gt}, is computed as either the Stone logarithmic price index

$$\pi_{gt} = \sum_{k \in g} s_{kt} \log p_{kt},$$

or the Deaton and Muellbauer exact price index

$$\pi_{gt} = \alpha_0 + \sum_{k \in g} \alpha_k p_k + \frac{1}{2} \sum_{j \in g} \sum_{k \in g} \gamma_{kj} \log p_k \log p_j.$$

The exact form of the price index does not seem to be very important for the results. If the latter is used the estimation is non-linear, while with the Stone index the estimation can be performed using linear methods.

Middle level: log-log demand:

$$\log q_{gt} = \beta_g \log Y_{Rt} + \sum_{k=1}^{G} \delta_k \log \pi_{kt} + \alpha_g + \epsilon_{gt};$$

$$g = 1, ..., G,$$

where: q_{gt} is the quantity of the *g*th segment in market *t;*
Y_{Rt} is total category (e.g., cereal) expenditure;
π_{kt} are the segment price indexes;

Top level: demand for the category is

$$\log Q_t = \beta_0 + \beta_1 \log I_t + \beta_2 \log \Pi_t + Z_t \delta + \epsilon_t$$

where: Q_t is the overall consumption of cereal in market *t*;
I_t is real income;
Π_t is the price index for the category;
Z_t are variables that shift demand;

Estimation: From lowest level up;
Linear or Non-linear (depending on price index);
Usually impose cross-equation restrictions;

Advantages: (1) Closely related to neo-classical demand;
(2) Flexible within a segment;

Disadvantages: (1) Requires division into segments;
(2) The above example does not exactly satisfy theory;
(3) Does not deal with: multicollinearity/IV or with heterogeneity;

Spatial Models

Combine neo-classical demand with a characteristics approach (Slade, UK brewing, Conley in various papers).

Suppose demand is

$$q = A + Bp + \epsilon,$$

where: q is a $J\times 1$ vector of quantities;
p is a $J\times 1$ vector of prices of the different products;
ε is a $J\times 1$ vector of error-terms;
A is a $J\times 1$ vector of parameters;
B is a $J\times$J matrix of parameters;

If J is large, then too many parameters to estimate.

Their solution:

$A_j = \alpha + X_j\beta,$ where X_j are the characteristics of product j.

$B = \{b_{ij}\}_{i,j=1\ldots J}$

where: b_{ii} = parametric function of X_i;
b_{ij} = non-parametric function of $X_i - X_j$;

Disadvantages: (1) The estimated functions are not promised to satisfy the requirements of economic theory. Cannot use them for welfare.

(2) Does not deal with: multicollinearity/IV or with heterogeneity;

(3) Results are not very promising;

Discrete Choice Models

Essentially solve the dimensionality problem by projecting the products onto a characteristics space, making the relevant dimension the dimension of the characteristics.

Let the (conditional) indirect utility of consumer *i* from product *j* in market *t* be

$$u_{ijt} = \alpha_i(y_i - p_{jt}) + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt},$$

$$i=1,\ldots,I_t \quad j=1,\ldots,J, \quad t=1,\ldots,T$$

where:

- y_i is the income of consumer *i;*
- p_{jt} is the price of product *j* in market *t;*
- x_{jt} is a $1\times K$ vector of observable characteristics of product *j;*
- ξ_{jt} is an unobserved (by the econometrician) product characteristic;
- ε_{ijt} is a mean zero stochastic term;
- α_i is consumer's *i* marginal utility from income;
- β_i is $K\times 1$ vector of individual specific taste-coefficients.

Unobserved characteristics can include the impact of unobserved promotional activity, unquantifiable factors ("brand equity") or systematic shocks to demand. We can model $\xi_{jt} = \xi_j + \xi_t + \Delta\xi_{jt}$ and capture ξ_j and ξ_t by brand- and market-specific dummy variables.

Implicit in the specification:
(1) Utility is quasi-linear; alternatives C-D $\log(y_i - p_{jt})$, or $f(y_i - p_{jt})$;
(2) All consumers value the unobserved characteristic in the same way;
(3) All consumers face the same product characteristics (prices)

The consumer-level taste parameters are modeled

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i \ , \quad v_i \sim P_v^*(v), \quad D_i \sim \hat{P}_D^*(D), \tag{5}$$

where: D_i is a $d\times 1$ vector of "observed" demographic variables;
v_i unobserved consumer attributes;
$P_v^*(\cdot)$ is a parametric distribution;
$\hat{P}_D^*(\cdot)$ is either a non-parametric distribution known from other data sources or a parametric distribution with the parameters estimated elsewhere;
Π is a $(K+1)\times d$ matrix of coefficients;
Σ is a $(K+1)\times(K+1)$ matrix of parameters.

The specification of the demand system is completed with the introduction of an "outside good": the consumers may decide not to purchase any of the brands. The indirect utility from this outside option is

$$u_{i0t} = \alpha_i y_i + \xi_{0t} + \pi_0 D_i + \sigma_0 v_{i0} + \epsilon_{i0t} \ .$$

Normalize ξ_{0t}, π_0 and σ_0 to zero;

Let $\theta = (\theta_1, \theta_2)$.
$\theta_1=(\alpha, \beta)$ contains the "linear" parameters;
$\theta_2=(\Pi, \Sigma,)$ the "non-linear" parameters.

Combining the equations

$$u_{ijt} = \alpha_i y_i + \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \mu_{ijt}(x_{jt}, p_{jt}, v_i, D_i; \theta_2) + \epsilon_{ijt}$$

$$\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}, \quad \mu_{ijt} = [-p_{jt}, x_{jt}] * (\Pi D_i + \Sigma v_i)$$

Consumers are assumed to purchase one unit of the good that gives the highest utility.

Define:

$$A_{jt}(x_{.t}, p_{.t}, \delta_{.t}; \theta_2) = \left\{ (D_i, v_i, \epsilon_{i0t}, \ldots, \epsilon_{iJt}) \mid u_{ijt} \geq u_{ilt} \quad \forall l = 0, 1, \ldots, J \right\}$$

where $x_{.t} = (x_{1t}, \ldots, x_{Jt})'$, $p_{.t} = (p_{1t}, \ldots, p_{Jt})'$ and $\delta_{.t} = (\delta_{1t}, \ldots, \delta_{Jt})'$ are observed characteristics, prices and mean utility of all brands, respectively.

Then, assuming ties occur with zero probability, the market share of the jth product is given by

$$s_{jt}(x_{.t}, p_{.t}, \delta_{.t}; \theta_2) = \int_{A_{jt}} dP^*(D, v, \epsilon) = \int_{A_{jt}} dP^*(\epsilon \mid D, v)\, dP^*(v \mid D)\, dP^*(D)$$

$$= \int_{A_{jt}} dP_{\epsilon}^*(\epsilon)\, dP_v^*(v)\, d\hat{P}_D^*(D) \;,$$

where $P^*(\cdot)$ denotes population distribution functions.

Given assumptions on the distribution of the (unobserved) individual attributes we can compute this integral.

Distributional Assumptions

(1) Logit:
Set $\theta_2 = 0$, which implies $\beta_i = \beta$ and $\alpha_i = \alpha$ for all *i,* and the indirect utility becomes

$$u_{ijt} = \alpha(y_i - p_{jt}) + x_{jt}\beta + \xi_{jt} + \epsilon_{ijt}.$$

Assume that ε_{ijt} are distributed iid according to a Type I extreme value distribution, then market share of brand j in market t, is

$$s_{jt} = \frac{\exp(x_{jt}\beta - \alpha p_{jt} + \xi_{jt})}{1 + \sum_{k=1}^{J} \exp(x_{kt}\beta - \alpha p_{kt} + \xi_{kt})} \;.$$

Appealing, due to its tractability, it restricts the substitution patterns. The price elasticities are

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} -\alpha p_{jt}(1 - s_{jt}), & \text{if } j = k; \\ \alpha p_{kt} s_{kt}, & \text{otherwise.} \end{cases}$$

own-price elasticities roughly proportional to price (heterogeneity);
cross-price elasticities do not depend on *j* (i.i.d. shocks);

(2) Nested Logit:

Divide the products into nests;
Set $\theta_2 = 0$, but allow for error components in ε_{ijt}, so products in the same nest have correlated errors;

This is equivalent to allowing a nest dummy variable as a characteristic and setting a particular distribution of its coefficient;

Issues: Does not deal with own-price elasticities;
Requires a-priori segments;
Order of "nesting" matters;

Principles of Differentiation Generalized Extreme Value model (Bresnahan, Stern and Trajtenberg) deals with this.

(3) Random Coefficients Logit/Mixed Logit :

Assume that ε_{ijt} are distributed iid according to a Type I extreme value distribution (can also allow for nesting);

Generate correlation through μ_{ijt}

The price elasticities are

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt})\, d\hat{P}_D^*(D)\, dP_v^*(v), & \text{if } j = k; \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt}\, d\hat{P}_D^*(D)\, dP_v^*(v), & \text{otherwise.} \end{cases}$$

where $s_{ijt} = \exp\{\delta_{jt} + \mu_{ijt}\} / (1 + \sum_{k=1}^{K} \exp\{\delta_{kt} + \mu_{ikt}\})$, is the probability of individual i purchasing product j.

Patterns in own-price elas driven by heterogeneity;

Patterns in cross -price elas driven by correlation in taste for characteristics: shocks for similar products will be correlated,

In principle, this model is flexible enough to approximate any choice model.

The cost is in the computation complexity;

Advantages (of discrete choice):

(1) Deals with all the problems we raised;

(2) Relatively easy data requirements;

Problems:

(1) What if choice is not discrete?

conceptual – define choice accordingly;

data – can be an issue if we observe consumer-level data;

questions – e.g., bundling;

(2) Restrictive function form/too few parameters;

(3) Identification;

Extensions:

Multiple Discreteness: Hendel (*RES*, 99) models the demand for computers by firms. Each firm buys several brands and several units.

Models: (1) firms tasks – multiple brands;

(2) decreasing MU from quantity – several units;

Discrete-Continuous choice: Dubin-McFadden (*EMA,* 84) model the choice of appliance + amount of usage;