Chapter 4:  
Product Differentiation *  

Steven Berry  
Yale Univ.  

Ariel Pakes  
Harvard Univ.  

October 20, 2003  

While homogenous goods is a convenient assumption for many models, it is frequently violated in practice. Many markets feature goods that are not identical, they vary in quality, features, reliability, reputation and/or geographic location. Indeed, markets of literally identical goods seem to be relatively rare, especially once differences in seller’s locations and reputations are taken into account.

Markets with differentiated products raise a number of important issues. Demand analysis obviously becomes much more complicated, as we can no longer think of a single market demand curve, but one rather must think of a demand system that yields the demands for each product in the market. It becomes even more plausible to think of the market as an oligopoly, rather than perfectly competitive, as the producers of each product will typically face downward sloping demand curves even in the presence of competition from other firms. Also, at some point (a topic considered in later chapters), we would like to know how it is that products come to be differentiated — what economic process generates new products with differing characteristics?

We begin this chapter with a review of simple theoretical models, leaving details to other sources, especially to Tirole.

*This is very much a work-in-progress. Corrections Welcome. Copyright by Steven Berry and Ariel Pakes.
1 Review of Theory.

Here we discuss various approaches to demand, and then discuss the static oligopoly equilibrium.

1.1 Deriving Demand.

The first question is from where does the demand system come from. There are two traditions in deriving discrete choice. The first tradition considers a representative consumer who has a taste for consuming a variety of products. Typically, the demand curve is derived from a well-specified utility function that features decreasing marginal utility from the consumption of each good in the market. This provides the incentive for the representative consumer to spread consumption across a variety of goods. The reason that goods are differentiated is typically buried in the parameters of the utility function, rather than being made explicit.

A second tradition locates goods in a space of product characteristics. Consumers have heterogeneous tastes and place differing utility weights on the different product characteristics. Typically, each consumer is modeled as buying at most one unit of the good, yielding a “discrete choice” model of demand at the consumer level. Product variety is then a response to the variety of consumers preferences, rather the the “taste for variety” of a representative consumer. Aggregate (market level) demand is then found by summing up the demands of the individual consumers.

At the market level, the differences between the two approaches can be exaggerated, as some aggregate demand curves can be derived from both approaches. Thus, at the market level the question is to some degree simply one of functional form and a question of which modelling approach gives better intuition for how that functional form restricts the eventual answers.

One advantage of the traditional characteristics approach is that it is easy to incorporate information about household-level demand. The representative consumer approach would face the problem of explaining why different households make different choices. However, there is no reason why the households level data must be modelled according to a discrete choice framework, as indicated by the discussion of Jorgensen style translog demand functions in an earlier section.

The choice between discrete and continuous models of demand will typically be decided by the nature of the market at hand. In many cases in
would be useful to combine discrete and continuous models (Hendel, cite.)

1.2 Representative Consumer Models.
On the theory side, we would discuss Dixit-Stiglitz style models (see Tirole),

1.3 Empirical Models, Non-Address Demand
These are models that typically assume a representative consumer, but always model demand as fundamentally for the product, not the characteristics of the product.

Separable demand and AIDS

In partial equilibrium demand analysis, there is an important question of whether we can really break out a subgroup of products and estimate demand for that subgroup. This is appropriate if the utility function is weakly separable. For example, we might group 4 products into 2 groups and consider the utility function

\[ u(q_1, q_2, q_3, q_4) = u(v_a(q_1, q_2), v_b(q_3, q_4)). \]

In this case, it is easy to show that demands for the goods in group a are a function only of the two prices \( p_1 \) and \( p_2 \) and the total group expenditure on \( x_a = p_1q_1 + p_2q_2. \) The within group expenditure is probably econometrically endogenous.

We then face the problem of modelling group expenditure. It would be nice if there were price indices \( \pi_a \) and \( \pi_b \) such that we could think of solving the “group expenditure” problem

\[
\max_{v_a, v_b} u(v_a, v_b) \text{ s.t. } \pi_a v_a - \pi_b v_b = Y.
\]

We could then first estimate within group demands, conditional on group expenditure for detailed data on a small number of products and then estimate the group expenditures from more aggregate expenditure.

This leaves open the question of what is the price index for the aggregate good. In special cases, known as “Gorman Polar Forms” (and “Generalized Gorman Polar Forms”) it is possible to derive theoretically correct aggregate price indices. In empirical practice, though, a CPI-style price index is usually
substituted for the theoretically correct price index. This is true in part because the researcher may not have access to all the individual prices in the other groups and in part because the parameters of the theoretical price index may enter the demand functions in a non-linear way. The problems of non-linearity are lessened with the advance of econometrics and of computer speed, but the problems of data availability are not.

Also, there is still the question of how to model demands at the lowest budgeting level. Deaton and Muellbauer suggest a particular group expenditure function that is a Gorman Polar Form and which generates the “Almost Ideal Demand System”. In particular, they assume group expenditure functions of

$$\log[e(u, p)] = a(p) + ub(p),$$

where

$$a(p) = \alpha_0 + \sum_k \alpha_k p_k + \frac{1}{2}\sum_{kl} \gamma_{kl}^* \log(p_k) \log(p_l),$$

and

$$b(p) = \beta_0 \Pi p^\beta_k.$$  

If we now solve for the share equations, which are typically the estimating equations, we get

$$w_i = \alpha_i + \sum_j \gamma_{ij} \log[p_j] + \beta_l \log[x/P],$$

where the theoretical price index $P$ is

$$\log[P] = \alpha_0 + \sum_k \alpha_k \log[p_k] + \frac{1}{2}\sum_i \sum_j \log[p_i] \log[p_j],$$

and the $\gamma$ parameters are given by $\gamma_{ij} = (1/2)\gamma_{ji}^* + \gamma_{ij}^* = \gamma_{ji}$. The group expenditure, $x$, is undoubtedly endogenous. As with cost functions, the theoretical restrictions on household demand can be tested via the estimated parameters.

**Functional Form and Household Level Demand: Translog Example**

The AIDS demand system is often applied to aggregate (market level) data where the researchers wants to estimate a demand system for many different categories of products. In this case, the nested structure of multi-stage budgeting helps greatly. However, the multi-stage budgeting comes at the expense of a very restrictive functional form. In other contexts one might
prefer to give up multi-stage budgeting and specify a more flexible demand system. For example, Jorgenson and Lau (1975) suggest a system for demand estimation that begins with a trans-log indirect utility function:

\[
\ln(V(p, Y_i, z_i, \epsilon_i)) = \text{constant} + \alpha' \ln(p/Y_i) + \frac{1}{2} \ln(p/Y_i)'B\ln(p/Y_i) + \ln(p/Y_i)'Tz_i + \ln(p/Y_i)'\epsilon_i,
\]

where

- \( \ln(p/Y_i) \) is a vector of prices normalized by income (to enforce h.d.0 in \( p \) and \( Y_i \)),
- \( z_i \) is a \( K \) by 1 vector of observed (by the econometrician) household attributes,
- \( \epsilon_i \) is a \( J \) by 1 vector of unobserved (by the econometrician) household attributes.
- the vector \( \alpha \) and the matrices \( B \) and \( \Gamma \) are parameters to be estimated.

\[
s_{ij} = \frac{\alpha_j + B_j \ln(p/Y_i) + \gamma_j z_i + \epsilon_{ij}}{\sum_r (\alpha_r + B_r \ln(p/Y_i) + \gamma_r z_i + \epsilon_{ir})},
\]

where \( B_j \) and \( \Gamma_j \) are the \( j^{th} \) row of the appropriate matrices. Jorgenson makes assumptions so that the denominator does not depend on household attributes, which generates a linear share equation. But computers are faster now and so there is no longer reason to insist on linearity. However, it is still necessary to put some restrictions on the model and traditional assumptions include \( \sum_r \epsilon_{ir} = 0 \) and \( \sum_r \alpha_r = 1 \).

1.4 Discrete Choice Models.

Many of the product differentiation models we consider are of the discrete-choice form. An advantage of these models is that they build demand from a well-specified utility for the characteristics of products. An unfortunate restriction is that they usually constrain each consumer to consider buy at most one unit of a good. This restriction can be relaxed: e.g. Hendel’s 1994 Harvard Ph.D. thesis. Examples of discrete choice models used in IO theory include the Hotelling and vertical models. Anderson, DePalma, and Thisse
(1992) provide a lengthy discussion of discrete choice models as used in the theory of product differentiation.

A general discrete choice model starts with by specifying the utility of consumer \( i \) for product \( j \) as

\[
u_{ij} = U(x_j, p_j, \nu_i), \]

where \( x_j \) is a vector of product characteristics, \( p_j \) is the price of the product and \( \nu_i \) is a vector of consumer characteristics. (Rather than model utility directly as a function of price, it might be preferable to model it as a function of expenditures on other products and then derive the “indirect” utility \( U \) as a function of price.)

As example, the utility function in the Hotelling model with quadratic transportation costs is

\[
u_{ij} = \bar{u} - p_j - (x_j - \nu_i)^2, \]

where \( x_j \) is the location of the product along the line and \( \nu_i \) is the location of the consumer. [brief derivations of Hotelling demand function here: note that demand is the integral over a region of the \( \nu \)'s.]

In the vertical model (e.g. (Shaked and Sutton 1982), (Bresnahan 1987)).

\[
u_{ij} = \nu_i x_j - p_j\]

where \( x_j \) is “quality”, and \( \nu_i \) is the consumers “taste” for quality. [Brief derivation of demand here. Note that as in Hotelling, a small variation in price or quality affects only the firms two “neighbors”. This is a result of the one-dimensional product space.]

Such utility functions can be extended to multiple characteristics. For example, an analog of the Hotelling model, with \( K \) characteristics, is

\[
u_{ij} = \bar{u} - p_j - \sum_{k=1}^{K} \alpha_k (x_{jk} - \nu_{ik})^2, \]

while an extension of the vertical model is the “pure random coefficients model”:

\[
u_{ij} = \left[ \sum_{k=1}^{K} \nu_{ik} x_{jk} \right] - p_j\]

A traditional econometric specification is:

\[
u_{ij} = x_j \beta_i - \alpha_i p_j + \epsilon_{ij}\]
where the consumer “tastes” are given by

\[ \nu_i = (\beta_i, \alpha_i, \epsilon_{i1}, \epsilon_{i2}, \ldots, \epsilon_{iJ}) \]

In the traditional econometric specification, the \( \epsilon_{ij} \) are assumed to be i.i.d. across products and consumers. In the simplest case, there are no random coefficients on the product characteristics, \( (\beta_i = \beta \text{ and } \alpha_i = \alpha) \) and the \( \epsilon_{ij} \) have the “type 2 extreme value” distribution

\[ F(\epsilon) = e^{-e^{-\epsilon}}. \]

This gives the traditional logit model, where the probability that good \( j \) is purchased (i.e. the market share of product \( j \)) is

\[ s_j = \frac{e^{\delta_j}}{\sum_{r=1}^{J} e^{\delta_r}}, \]

with \( \delta_j = x_j \beta - \alpha p_j \). If we add an “outside good” with \( \delta_j = 0 \), the logit market share becomes:

\[ s_j = \frac{e^{\delta_j}}{1 + \sum_{r=1}^{J} e^{\delta_r}}. \]

While the logit market share is easy to calculate, the model has unintuitive properties. In particular, cross-price effects do not depend on the degree to which the products have similar \( x_j \)'s, but only on the values of the sum \( \delta_j \).

The counter-intuitive substitution patterns do not come only from the specific distributional assumption in the logit model, but from the assumption that the only variance in consumer tastes comes through the i.i.d. product-specific terms \( \epsilon_{ij} \). Since these terms are i.i.d., there is no source of correlation in consumer tastes across similar products.

When variance is added to the term \( \beta_i \) and/or \( \alpha_i \), then substitution patterns can become more reasonable. Now, a consumer who buys a good with a large value of some characteristic is more likely than the average consumer to have as a second choice another good with a large value of that characteristic.

However, as long as the distribution of \( \epsilon \) is i.i.d., there are effectively as many product characteristics as there are products. Indeed, we can think of \( \epsilon_{ij} \) as the consumer taste for a product characteristic that is defined to be equal to one for product \( j \) and zero otherwise. If the \( \epsilon \)'s have an unbounded distribution, then all goods are strict substitutes for one another.
(e.g. $\partial s_j / \partial p_r$ is positive for all $j$ and $r$.) Contrast this to the one-dimensional taste models (Hotelling, vertical) in which each product is a substitute with only 2 other “nearby” products.

In a 1991 paper, Caplin and Nalebuff (1991) consider a broad class of discrete-choice models and provide a result on the existence of equilibrium in such models. They also provide a partial characterization of the equilibrium. They require two assumptions on the distribution of consumer utility. The first requires utility to be linear in consumer characteristics; note that all of our examples above satisfy this requires (although the quadratic transport cost example requires some re-writing to see this.) The second assumption places a restriction on the density of consumer tastes. We consider these assumptions in more detail in class.

1.5 Equilibrium

Having discussed the various methods of deriving demand, we can turn to a consideration of equilibrium. The simplest models of product differentiation would consider a set of single product firms each producing a differentiated product. We could begin by specifying a demand system for this set of related products, together with cost functions and an equilibrium notion. The usual assumption is Nash in prices. To analyze the case of equilibrium with differentiable demand, note that the profits of firm $j$ are given by

$$\pi_j(p) = p_j q_j(p) - C_j(q_j(p)).$$

The first-order condition is:

$$q_j + (p_j - mc_j) \frac{\partial q_j}{\partial p_j} = 0.$$

Note that we can rewrite this as

$$p_j = mc_j + b_j(p),$$

where the markup is

$$b_j(p) = \frac{q_j}{\left| \frac{\partial q_j}{\partial p_j} \right|}.$$

Also, we can write the product Lerner index in terms of the usual “inverse elasticity” rule.

$$\frac{(p_j - mc_j)}{p_j} = 1/\eta_j,$$

8
where $\eta_j$ is the absolute value of the product-specific elasticity.

The second order-condition is given by:

$$\frac{\partial q_j}{\partial p_j} + (p_j - mc_j) \frac{\partial^2 q_j}{\partial p_j^2} - \frac{\partial mc_j}{\partial q_j} \left( \frac{\partial q_j}{\partial p_j} \right)^2 < 0$$

This is surely negative when demand slopes down and is concave and marginal cost does not slope down. If demand is too convex, we sometimes have trouble (e.g. remember the monopoly case with constant elasticity demand, with an elasticity less than 1.) If the second-order condition holds everywhere, we have a sufficient (but not necessary) condition for a well-defined reaction function.

Differentiating the first-order condition, we find that prices are strategic complements when

$$\frac{\partial q_j}{\partial p_k} + (p_j - mc_j) \frac{\partial^2 q_j}{\partial p_j \partial p_k} - \frac{\partial mc_j}{\partial q_j} \frac{\partial q_j}{\partial p_k} \frac{\partial q_j}{\partial p_j} > 0 \quad (3)$$

This obviously holds in the linear demand case when the goods are substitutes, but need not hold in a non-linear case. This accounts for the general intuition that prices are strategic complements in the price-setting model.

### 1.5.1 Non-differentiable Demand

Some care must be taken in assuming differentiable demand functions, because many natural specifications result in discontinuous demand functions. For example, consider the simple Hotelling model of competition on a line, with the transportation cost equal to the absolute value of the distance between the consumer and the firm

$$u_{ij} = \bar{u} - p_j - \|x_j - \nu_i\|,$$

In this case it is possible for one firm to undercut ... [AND SO FORTH]

### 1.5.2 Quantity vs. Price Setting

In differentiated products models, one could also consider quantity setting firms, as in the Cournot model. In this case, the first-order condition becomes

$$p_j + q_j \frac{\partial p_j}{\partial q_j} - mc_j = 0.$$
If
\[ \frac{\partial p_j}{\partial q_j} = \frac{1}{(\partial q_j/\partial p_j)}, \]
(4)
then this first-order condition is the same as the price-setting f.o.c. Condition (4) holds when a change in price changes only the own-product quantity: this is just the monopoly case. However, in general, \( \partial p_j/\partial q_j \) is the \( j^{th} \) diagonal element of the \( J \) by \( J \) matrix
\[
\left[ \frac{\partial q}{\partial p} \right]^{-1},
\]
which is not the same as the inverse of \( \partial q_j/\partial p_j \). For many examples, the quantity-setting markup will be higher than the price setting markup.

The question of price vs. quantity setting in differentiated product models is asked, in a Kreps-Scheinkman like way, in J. Friedman (RAND, 1988) and I. Hendel (1994).

**Exercise 1** Exercises on simple models of Product Differentiation

1. Give sufficient conditions for each term of (3) to be positive.

2. Show that the derivatives of market share in the logit model are \( \partial s_j/\partial p_j = -\alpha s_j(1 - s_j) \) and \( \partial s_j/\partial p_r = \alpha s_j s_r \).

2 Empirical Work.

2.1 Representative Consumers Models

While we will mostly discuss empirical applications that involve discrete choice models, we begin with Hausman (1997), a nice application of AIDS models to a classical problem of differentiated products markets: how to evaluate new goods. This follows directly from the theory of demand for a representative agent. We go to the reference period; the period in which the new good is purchased (it is not purchased in the base period). We estimated demand systems in that period
\[
q_n = q_n(p_1, \ldots, p_{n-1}, p_n, y).
\]
(5)

We can use that demand system to compute the demands were all prices but the new goods price kept the same but the new goods price was raised
until just that value at which demands were cut off. That price will be called the virtual price, \( p^* \), (the idea of a virtual price is discussed by Hicks and formalized by Rothbarth).

Next use the estimated demand curves to solve for the expenditure function for a given level of utility, i.e. substitute \( q(\cdot) \) into \( u(\cdot) \) and invert to get
\[
x = e(p_1, \ldots, p_n, u^1)
\]
(6)

Calculate also
\[
x^* = e(p_1, \ldots, p_{n-1}; p_n^*, u^1).
\]
(7)

Hausman calls the exact COL index \( P(p, p^*, u^1) \equiv x^*/x \). This is a comparison of the utility in the second period were the new good not introduced to the utility when the good is introduced. It implicitly assumes that the old goods prices would have remained unchanged were the new good not introduced; because of oligopolistic markets and multiproduct firms this would not be so. This brings him to his analysis of “imperfect competition”.

I would have argued that all this is besides the point. What you want is to use the base period utility and prices to define the denominator of the index. Then
\[
P(p^0, p^1, u^0) = \frac{e(p_1^1, \ldots, p_n^1; p_n^*, u^0)}{e(p_1^0, \ldots, p_{n-1}^0, p_n^*, u^0)}.
\]
(8)

This \( P \) is the exact COL index; and does nto depend on the nature of competition.

Note. This whole approach requires estimation of the demand functions in the second period, in order to get demands as a function also of the new goods. On the other hand the traditional index number procedure (which does not compare new goods) does not handle new goods at all.

\[\textbf{2.1.1 Ready to Eat Breakfast Cereal Industry.}\]

This is an industry of five or six multiproduct firms, where the products turnover quite quickly. Between 1980-92, 190 new brands were introduced (160 existed at the start of the period). By 1992 95, or about half, of the new brands had been discontinued. Those new brands that remained controlled about a quarter of the market. Brands themselves all have a small market share, although the brand-combined shares of the manufacturers can be large. Aside from store brands, or private labels, there has been no significant entry in this market in over fifty years.
Data. Weekly cash register data from stores in major metropolitan areas for a two year period. Quantity variable is pounds of cereal of each brand sold.

Estimates a 3-level multistage budgeting system. The top level gives the demand for cereals. The mid level is split into market segments (kids, adults, and family). The lower level does brand choice. The lower level is estimated first, and uses the AIDS demand system, i.e.

\[ w_{int} = \alpha_{in} + \beta_i \log\left(\frac{x_{Gnt}}{P_{nt}}\right) + \sigma_j \log[p_{jn}] + \epsilon_{int}, \quad (9) \]

where \( i \) gives brands, and \( n \) is city. Note that the shares are taken for the city, and the expenditure is, as it should be, segment level expenditure, divided by a segment price index. He doesn’t tell us what price index he uses here.

The estimates of the lowest level imply a price index for each segment (at least for a given level of utility). Using those indices he uses a mid level demand system

\[ \log[q_{mnt}] = \beta_m \log[x_Bnt] + \sigma_k \delta_k \log[\pi_knt] + \alpha_{mn} + \epsilon_{mnt}, \quad (10) \]

where the first index \( m \) or \( k \) gives us the segment. Here \( x_B \) is total cereal expenditures.

Note that there is a restriction on this function across equations. The additivity implies that \( q_{mnt} \) is a function of \( \lambda(b_1, \ldots, b_G, x) \times b_G \). Thus the coefficients of other \( b \)'s are constrained across equations.

The top level is given by

\[ \log[u_t] = \beta_0 + \beta_1 x_t + \beta_2 \log[\Pi_t] + Z_t \delta + \epsilon_t, \quad (11) \]

where \( u_t \) is overall consumption of cereal, \( x_t \) is deflated disposable income, \( \Pi_t \) is a price index for cereals. This upper level demand system is estimated using aggregate monthly level data (sixteen years of it), and instruments (factors that shift costs, like costs of ingredients, packaging, and labor).

Given this system there is a question of consistent estimation of the mid and lower level systems. The best ways to estimate this depends on where the errors are generated from. There are two familiar reasons for generating an error in demand analysis;

- Variance in demand parameters not captured by the model. In a broader sense we usually consider these parameters to be a function of
perceptions of the product (which may vary over time, or across cities), and with characteristics of the population (family composition, types of stores, etc.). These differences cause differences in demand given price by region, period, etc. On the other hand, producers are setting prices with these differences in demand in mind, so we would expect prices to be correlated with the error. Consequently least squares estimation is biased.

- Errors in measurement in prices. Here again least squares is biased.

To construct instruments he uses

\[
\log[p_{jnt}] = \delta_j \log[c_{jt}] + \alpha_{jn} + w_{jnt},
\]

The predicted values from this equation are essentially the instruments. Identification is from the way costs moves over time within cities.

See Estimates.

Calculates a virtual price of about $7 with a standard error of $1.33, and the lower bound of a confidence interval at about $4.75 – 35% higher than the average price of apple-cinnamon cheerios. This translates into a $78 mill. increase in consumer surplus. If this was typical of brand introductions for this industry it would cause the price index for cereals to be overestimated by about 25%.

2.2 Estimating Discrete Choice Models on Market Level Data

We consider two discrete-choice papers in detail (Bresnahan 1987) and (Berry, Levinsohn, and Pakes 1995), both of which consider an application to the automobile industry.

2.3 The 1955 Auto Price War

(Bresnahan 1987) starts with a puzzle, which is the apparent decrease in automobile prices that occurred in 1955, which was an economic boom year, as compared to the surrounding years of 1954 and 1956, when automobile prices seemed to be higher. Bresnahan wanted to know whether differences in competition could account for the price change. One hypothesis is that
collusive behavior collapsed in the face of the economics boom. The idea, then, as in the homogeneous goods literature on “testing the competition”, is to look for evidence on what sort of competitive regime best explains the data in each of years 1954-1956.

Bresnahan assumes that we observe, for each product (= car), a set of product characteristics $x_j$ as well as the market outcomes of price and quantities sold, $p_j$, $q_j$. We observe the data only at the overall market (aggregate) level, without any information on the actions of individual households. Such data is readily available from industry-oriented publications (such as Automotive News and Ward’s.)

Bresnahan assumes that demand is given by the vertical model of the prior section. This implies that there is only one attribute per product, whereas the data (and intuition) suggest a number of different attributes such as size, horsepower and luxury. Bresnahan creates an index of these separate attributes, say $\delta_j = x_j \beta$, where $\beta$ is a parameter to be estimated. As opposed to more complicated spaces of product characteristics, the vertical model is used for analytical convenience. Demand is then given a function of qualities and prices of all the products:

$$q_j = q_j(\delta_j, \delta_{-j}, p_j, p_{-j}).$$  \hspace{1cm} (13)

(derive this.)

The prices are endogenously determined by a price-setting Nash equilibrium. Note that in the automobile industry, the first-order condition has to be modified to account for multi-product firms. GM, for example, offers dozens of products in the market today. To solve for equilibrium prices, one must also specify a cost side of the model. Costs, particularly marginal costs, are not typically observed so, as in the Cournot models above, we often are forced to infer marginal costs with the help of the oligopoly first order condition. Bresnahan assumes that marginal costs are a convex function of quality [question for the reader: why convex?].

This allows him to [FILL THIS OUT]:

- Solve for reduced-form $p$ and $q$.
- Tack on errors (say they’re normal.)
- Estimate by MLE under two equilibrium assumptions: Nash in price and collusion.
Given the estimates under the two competing equilibrium assumptions, Bresnahn does a “non-nested” hypothesis tests and finds, as conjectured, that collusive pricing fits better in 1954 and 1956, but Nash pricing fits better in 1955.

2.4 Endogenous Prices and Aggregation in Discrete Choice Models

In this section we discuss the method of Berry (1994) and Berry, Levinsohn, and Pakes (1995). The idea is to extend supply and demand empirical models to markets with product differentiation. As in the supply and demand literature, want to account for the implications of consumer and firm optimization together with equilibrium pricing.

Data are at least market-level observations on prices, quantities and characteristics of products. Might also have data on characteristics of purchasing consumers and on firm production (but this is rare.) Consumers modeled in a discrete choice framework. The utility of consumer $i$ for product $j$ depends on observed and unobserved (by us) characteristics, $(y, \nu)$, of the consumer and observed and unobserved (by us) characteristics of the product, $(x, \xi)$. Each firm produces a given set of products whose production cost depends on a vector of cost shifters $w$; presumably, $w$ includes the product characteristics, $x$. The equilibrium is Nash in prices.

To derive demand, we assume that choices are all products in the market and an outside good. The utility of a choice is determined by a parametric form for the interaction between consumer characteristics and product attributes (Lancaster, McFadden, etc.) The demand function is then derived by explicitly aggregating over the choices of consumers with different characteristics

2.4.1 Logit example.

In the simple logit case,

$$u_{ij} = x_j \beta - \alpha p_j + \epsilon_{ij} \equiv \delta_j + \epsilon_{ij},$$  \hspace{1cm} (14)

In logit model, with $u_{ij} = \delta_j + \epsilon_{ij}$, market share, $s_j(\delta)$, depends only on $\delta$, not on values of $x$ and $p$
Thus, the mean utility level, $\delta_j$, determines all behavior including market share and cross-price effects. Markups for price-setting firms vary only with market share.

Well-known solution: Interact Product and Consumer Characteristics. Random coefficients example:

$$u_{ij} = x_jp_i - \alpha j + \epsilon_{ij}. \quad (15)$$

More generally,

$$u_{ij} = \delta_j + \mu(x_j, p_j, \nu_i, \theta) + \epsilon_{ij}. \quad (16)$$

### 2.4.2 More on Functional Form

Problem with $\epsilon_{ij}$’s: effectively add new product characteristic every time add new product. Any product has positive sales and all products are substitutes. Extension to “no $\epsilon$” model is difficult but possible.

Vertical model gets rid of $\epsilon_{ij}$’s. But, all substitution is local – only with next highest and next lowest $p$. Substitution still doesn’t depend on $x$.

To get aggregate demand have to compute a complicated integral. Aggregation problem solved by simulation, as in Pakes (1986).

### 2.4.3 Endogenous Prices

As in traditional homogeneous goods models, the econometric endogeneity of prices follows from the presence of unobserved characteristics.

In autos unobserved characteristics include style, dealer quality, tradition. How to explain: Chevy Cavalier, Mazda Miata, BMW 300’s.

Assuming no unobserved product characteristic also leads to over fitting problem.

Unobserved characteristics are usually ignored in discrete choice literature. Micro studies sometimes add product-specific intercept, but often have too many choices. Also, product intercept is co-linear with $p, x$.

### 2.4.4 Solution

Allow for unobserved characteristic, $\xi_j$

$$u_{ij} = x_jp_i - \alpha j + \xi_j + \epsilon_{ij}, \quad \text{or} \quad (17)$$
Problem: $\xi$ enters $s_j$ in a highly non-linear fashion and is correlated with (at least) prices. Can solve for $\xi$ given parameters of model. However, one $\xi$ for every observation, need more restrictions. Following traditional demand literature, we assume that $\xi$ is uncorrelated with some “demand shifters”; here these are product characteristics. Could use other restrictions (endogenous characteristics?)

2.4.5 Inverting the Market Share Function

For the true model: $s_j = s_j(\delta, \theta)$.

Invert to find mean levels of utility:

$$\delta = s^{-1}(s, \theta) \quad (18)$$

Now could estimate demand parameters by IV from, e.g.,:

$$\delta_j = x_j \beta - \alpha p_j + \xi_j, \quad (19)$$

Alternatively, can think of solving for $\xi_j$ directly.

**Instruments** can come from excluded cost factors or characteristics of other firms.

2.4.6 Logit Example

Shares:

$$s_j(\delta) = e^{\delta_j} / (1 + \sum_r e^{\delta_r}), \quad (20)$$

Share of outside good:

$$s_0(\delta) = 1 / (1 + \sum_r e^{\delta_r}). \quad (21)$$

$$\Rightarrow \ln(s_j) - \ln(s_o) = \delta_j = x_j \beta - \alpha p_j + \xi_j. \quad (22)$$

Berry (1994) also gives analytic inverse for vertical model, nested logit. More complicated models require a numerical inverse.

In the logit case, estimate

$$y_j \equiv \ln(s_j) - \ln(s_o) = \delta_j = x_j \beta - \alpha p_j + \xi_j. \quad (23)$$

by 2SLS ($p_j$ is endogenous).
2.4.7 Firm Behavior

Marginal Cost:

\[ mc_j = w_j \gamma + \omega_j. \]  

(24)

Might also model \( \ln(mc_j) \).

Profits:

\[ \pi_f = \Sigma_{j \in J_f} (p_j - mc_j) M s_j(p, x, \xi; \theta), \]  

(25)

First-order Condition:

\[ s_j(p, x, \xi; \theta) - \Sigma_{r \in J_f} (p_r - mc_r) \frac{\partial s_r(p, x, \xi; \theta)}{\partial p_j} = 0. \]  

(26)

Given the demand function, it is possible to solve this for the vector of \( mc \)'s and so for markup, \( b_j(p, x, \xi, \theta) \) and for \( \omega_j \).

2.4.8 Estimating from the FOC

Given the demand parameters, can think of estimating the equation,

\[ mc_j = p_j - b_j(p, x, \xi, \theta) = w_j \gamma + \omega_j. \]  

(27)

Just as in estimating demand, estimates of the parameters \( \gamma \) can be obtained from orthogonality conditions between \( \omega \) and appropriate instruments. Can also estimate supply and demand together.

Notes:

- We do not require a unique equilibrium
- the markup depends on \( \xi, \omega \) and so is econometrically endogenous
- other static equilibria are easy (e.g. qty-setting, collusion)

Caveats:

- Nash Pricing
- no dynamics
- no production data
- no direct data on consumers
2.4.9 BLP

Question of interest: substitution patterns in auto demand; policy issues such as effect of Voluntary Export Restraints (VERs).

- Data: 20 years of public data on $p, x, q$ of approx 100 automobile models per year. Also use info on distribution of consumer income.

- Utility Function:

$$u_{ij} = x_j^\beta - \alpha_i p_j + \sum_k \sigma_k x_{jk} v_{ik} + \epsilon_{ij}.$$  \hspace{1cm} (28)

- Firm behavior: $\ln(mc) = w \gamma + \omega$, multi-product firms.

Econometric issues include simulating the shares, solving for $\xi$ and asymptotics in number of products and/or time.

References


