Differentiated Products Demand Systems (A)

Jonathan Levin

Economics 257
Stanford University

Fall 2009

Differentiated Products Demand - Outline

- Overview
- Supply side
- Product space
- Characteristic space
- Recent developments
- Class Discussion

Why do we care?

- Products in almost all markets are differentiated to some extent. Products differ in their physical characteristics, location, etc. This is true even for markets that are seemingly homogeneous (e.g. drugs, gasoline).
- We need some knowledge of the demand system in order to answer a broad set of questions. For example:
  - Measuring market power
  - Assessing the welfare effects of mergers
  - Assessing the effects of new taxes or tariffs
  - Analyzing product introduction
  - Estimating the welfare effects of new goods
- We may also be interested in consumer behavior per se: who buys what products, and the importance of search costs, switching costs, consumer information, advertising brand loyalty and so forth for purchasing behavior.

Some preliminary comments

- Theory on differentiated products (Anderson, de Palma, and Thisse; Shaked and Sutton; Caplin and Nalebuff; and others) is mainly focused on the supply side: given a demand system, what are the products and prices chosen in equilibrium?
- Due to various reasons (data, generality, and complexity) the empirical literature has tended to focus more on the demand side. We will discuss the simple Bertrand-Nash supply model used most often and spend the rest of the time on different alternatives to specify demand.
- Note that in many regulated industries (electricity, water, postal services) prices are set through a non-market mechanism, making the demand side the natural place to start.

Some preliminary comments

- The individual choice aspect of demand often leads to introspection and a tendency to make the models richer and "more realistic". We should never forget that there is a clear trade-off!
- Think about adding richness from a cost-benefit analysis: richer models are typically more complicated to compute, test, and to communicate to others. These costs should be compared to the added-value the complexity buys us. If this value is small, we may be better off with a simpler model.
- You will also see that many demand papers focus a lot on specifying a theoretically appealing econometric model, and often put less emphasis on whether the data have the relevant variation to identify the key parameters.
- Remember that empirical models should always be evaluated in light of the question at hand and the data available to answer it.

The supply side

- We begin with the supply side: it is often similar across applications and it helps to identify the parameters and issues we're going to care about.
- But wait ... why model supply if we are estimating demand?
- Typical assumption is that prices are the outcome of a Bertrand-Nash Equilibrium. (does this make sense? when?)

The supply side: monopoly

- Lerner condition relates the firm's mark-up to its price elasticity:

$$\frac{p-c}{p} = \frac{1}{\eta} = \frac{q(p)/p}{q'(p)}$$

- Infering costs: if we assume optimal pricing, knowledge of η allows us to learn c, ... or vice-versa. Alternatively, knowledge of η and c allows us to test for optimizing behavior.
- What is c anyway? (example: MS Windows).
- Suppose we want to know the impact of a cost shock. Note that:

$$p = c + \frac{1}{q'(p)/q(p)}.$$

Pass-through of cost changes depends on the curvature of demand $q(p)$.

The supply side: oligopoly

- More generally, each firm solves:

$$\max_{\{p_i\}_{i\in f}} \left\{ \sum_{i\in f} p_i q_i(p) - \sum_{i \in f} c_i(q_i(p)) \right\}$$

- This leads to FOCs (assume SOCs are satisfied):

$$\forall i \in f: \; q_i(p) + \sum_{j\in f} p_j \frac{\partial q_j(p)}{p_i} - \sum_{j\in f} mc_j(q_j) \frac{\partial q_j(p)}{p_i} = 0$$

- In matrix notation:

$$q + (\Omega \cdot D_p q)(p - mc) = 0$$

where Ω is an "ownership matrix".

The supply side (cont.)

- Under regularity conditions that guarantee invertibility of $(\Omega \cdot D_p q)$ we have:

  $$p = mc - (\Omega \cdot D_p q)^{-1} q$$

  i.e. price is equal to marginal cost plus a markup.
- This is the multi-firm, multi-product analogue of the Lerner condition.
- With the demand system specified and its parameters estimated, we can now:
  - Calculate markups and recover marginal costs (but we need instruments, i.e. cost shifters, to do so; we typically assume the error term enters additively into *mc*, which makes the econometrics simpler).
  - But, hey ... why not use data on costs?
  - Carry counterfactual analysis: mergers, benefits of new goods, etc. This is the main reason why we need a more structural model.

The supply side - applying the model

- Researchers often estimate their demand model without assuming that prices are eqm prices — although they typically justify the validity of their instruments by invoking some assumptions about price setting. Adding "supply side moments" can provide additional identifying power but of course requires additional assumptions — a trade-off.
- Since the NEIO revolution, there has been a reluctance to use reported cost data, particularly accounting data, leading many IO studies to favor cost estimates obtained from an eqm assumption. In many industries, however, costs can be measured fairly accurately. In a given application, you should ask yourself what approach makes sense.

Product vs. Characteristic space

- There are two general approaches to estimating demand. Product space is more natural: consumers have preferences over products, and those preferences lead to demand at the product level. The characteristic space approach (Lancaster, 1966; McFadden, 1973) views products as bundles of characteristics, and consumers having preferences over those characteristics, rather than over the bundles.
- Some benefits of the characteristic approach:
  - Reduce the number of demand parameters.
  - Evaluate demand for new goods (but not always, e.g. laptop).
  - Builds up consistently from individual decisions.
- Some drawbacks:
  - More data needed
  - Often leads to complex estimation methods
  - The "too many characteristics problem" (see later)
  - Unmeasurable attributes (e.g. books or movies).

Demand in product space

- The typical problem is to estimate $q = f(p, z)$ in a way which is flexible and consistent with theory (aggregation is the main issue).
- Models who do so:
  - Linear Expenditure Model (Stone, 1954)
  - Roterdam Model (Theil, 1965; Barten, 1966)
  - Translog Model (Christensen, Jorgensen, and Lau, 1975)
  - Almost Ideal Demand System (AIDS) (Deaton and Muellbauer, 1980)
- Problem in applying these models in practice:
  - Dimensionality: even a simple linear demand model $q = Ap$ would require the number of parameters to be in the order of J^2, which is often too much.
  - Consumer heterogeneity: the above methods estimate aggregate demand. Many economic questions may benefit from explicit modelling of consumer heterogeneity, particularly if we have data on individual decisions.
  - Identification: to estimate demand we need sufficient variation in prices to identify the parameters. Where will this come from?

Demand in product space - solutions

- For certain economic questions, we can skirt the problems by focusing on demand for a single product or class of products. Otherwise...
- Simple representative consumer models:
  1. The Constant Elasticity of Substitution (CES) model:

$$u(q) = \left(\sum q_i^{\rho}\right)^{1-\rho} \implies q_i = \frac{p_i^{-1/(1-\rho)}}{\sum p_j^{-\rho/(1-\rho)}} I$$

     Just one parameter! But now own- and cross-price elasticity is the same for all products. Seems implausible ... and potentially misleading (why? when?).
  2. Logit demand (Anderson et al., 1992)

$$u(q) = \sum \delta_i q_i - \sum q_i \ln q_i$$

     This has the sometimes problematic IIA property: elasticities depend only on market shares (the δ's in this case), but not on the similarities among products.

These models are useful, but only for a particular set of questions (e.g. in international trade they use them all the time). Typically, for

Separability, and multi-stage budgeting

- Key idea: solve the dimensionality problem by dividing products to small groups and sub-groups; allow flexible substitution within groups.
- To make this consistent with theory, we need two related assumptions: separability in preferences and multi-stage budgeting.
  - Separability: preferences for products of one group are independent of product-specific consumption of products from other groups, eg.

    $$U(q) = f\left(u_1(q^{(1)}), ..., u_n(q^{(n)})\right)$$

    where $\{q^{(j)}\}$ is a partition of q.
  - Multi-stage budgeting: consumers can allocate total expenditures in stages, to sub groups, to sub-sub groups, and so forth. At each stage, only the prices (or price indices) of members of the group matter.

These are similar but are not the same or nested.

Separability, and multi-stage budgeting

- For empirical purposes, the following two sufficient conditions are typically assumed:
  - Indirect utility for each segment is of the Generalized Gorman Polar Form (Gorman, 1959)
  - Overall utility is additively separable in the sub-utilities.

Almost Ideal Demand System (AIDS)

- Deaton & Muellbaur developed the AIDS model to study demand for broad classifications of products (food, housing, clothing).
- Hausman et al use it for studying differentiated product markets. They have three levels of aggregation: the whole market, market segments, and individual products at the lowest level.

AIDS and its applications

- Demand for product i in segment g is given by:

$$w_i = \alpha_i + \beta_i \log(y_g / P_g) + \sum_{j \in g} \gamma_{ij} \log p_j + \varepsilon_i$$

  - w_i is the within-segment expenditure share.
  - y_g is expenditure on segment g.
  - P_g is segment g's price index, p's are prices.
- Attractions: flexible, aggregates over individuals, easy to test or impose theoretical restrictions (e.g. symmetry of Slutsky matrix).
- Possible price indices include Stone's logarithmic index:

$$P_g = \sum_{j \in g} w_j \log p_j$$

  or Deaton and Muellbauer's exact price index

$$P_g = \alpha_0 + \sum_{j \in g} \alpha_j p_j + \frac{1}{2} \sum_{j \in g} \sum_{k \in g} \gamma_{jk} \log p_j \log p_k$$

  which is theoretically appealing but more complex for estimation.

AIDS and its applications (cont.)

- The middle level can be estimated by either AIDS again, or by a log-log specification. None is fully consistent with theory. AIDS would look the same (substitute price indices for prices), while log-log would be

$$\log q_g = \alpha_g + \beta_g \log y + \sum_h \delta_{g,h} \log P_h + \varepsilon_g$$

- The top level is a single log-log equation (consistent with theory), with the addition of a set of demand shifters Z:

$$\log q = \alpha + \beta \log y + \delta \log P + Z\gamma + \varepsilon$$

AIDS and its applications (cont.)

- Two important problems:
  - Underlying theory of individuals assumes no corner solutions (i.e. each consumer buys some of each good). While this is true for broad categories, it may not be true for specific products.
  - We need to a-priori classify the products, which is not always a clear cut (this will also show up later, with nested logit).
- We will go over two similar applications, for beer and cereal. Note that these two are careful choices, as here the segmentations is relatively clear (compared to, say, PC's or autos).

Hausman, Leonard, and Zona (1994)

- Data: beer transaction data across US cities aggregated to monthly prices and quantities (is this what we want?).
- Estimation using AIDS as above. Three segments: light beers, premium (e.g. Miller, Budweiser, Coors), and popular priced. Five brands in each. No supply side equation.
- Instruments: variation in prices of the same brand in other cities. The idea is that prices satisfy:

$$\log p_{jct} = \alpha_{jc} + \log c_{jt} + \omega_{jct}$$

  and the ω_{jct} are independent of each other.
- Does this make sense? What can make it a bad strategy? How would this bias the coefficients?
- Specification test for the segmentation: prices of products in other segments should not affect beyond their effect through the price index.

Hausman et al. Estimates

TABLE 1

Beer Segment Conditional Demand Equations.

	Premium	Popular	Light
Constant	0.501	−4.021	−1.183
	(0.283)	(0.560)	(0.377)
log (Beer Exp)	0.978	0.943	1.067
	(0.011)	(0.022)	(0.015)
log ($P_{PREMIUM}$)	−2.671	2.704	0.424
	(0.123)	(0.244)	(0.166)
log ($P_{POPULAR}$)	0.510	−2.707	0.747
	(0.097)	(0.193)	(0.127)
log (P_{LIGHT})	0.701	0.518	−2.424
	(0.070)	(0.140)	(0.092)
Time	−0.001	−0.000	0.002
	(0.000)	(0.001)	(0.000)
log (# of Stores)	−0.035	0.253	−0.176
	(0.016)	(0.034)	(0.023)

Number of Observations = 101.

Hausman et al. Estimates

Brand Share Equations: Premium.

	1 Budweiser	2 Molson	3 Labatts	4 Miller	5 Coors
Constant	0.393	0.377	0.230	−0.104	–
	(0.062)	(0.078)	(0.056)	(0.031)	–
Time	0.001	−0.000	0.001	0.000	–
	(0.000)	(0.000)	(0.000)	(0.000)	–
log (Y/P)	−0.004	−0.011	−0.006	0.017	–
	(0.006)	(0.007)	(0.005)	(0.003)	–
log ($P_{\text{Budweiser}}$)	−0.936	0.372	0.243	0.150	–
	(0.041)	(0.231)	(0.034)	(0.018)	–
log (P_{Molson})	0.372	−0.804	0.183	0.130	–
	(0.231)	(0.031)	(0.022)	(0.012)	–
log (P_{Labatts})	0.243	0.183	−0.588	0.028	–
	(0.034)	(0.022)	(0.044)	(0.019)	–
log (P_{Miller})	0.150	0.130	0.028	−0.377	–
	(0.018)	(0.012)	(0.019)	(0.017)	–
log (# of Stores)	−0.010	0.005	−0.036	0.022	–
	(0.009)	(0.012)	(0.008)	(0.005)	–
Conditional Own	−3.527	−5.049	−4.277	−4.201	−4.641
Price Elasticity	(0.113)	(0.152)	(0.245)	(0.147)	(0.203)

$$\Sigma = \left\{ \begin{array}{cccc} 0.000359 & -1.436E-05 & -0.000158 & -2.402E-05 \\ - & 0.000109 & -6.246E-05 & -1.847E-05 \\ - & - & 0.005487 & -0.000392 \\ - & - & - & 0.000492 \end{array} \right\}$$

Note: Symmetry imposed during estimation.

Hausman et al. Estimates

Light Segment Own and Cross Elasticities.

	Genesee Light	Coors Light	Old Milwaukee Light	Lite	Molson Light
Genesee Light	−3.763	0.464	0.397	0.254	0.201
	(0.072)	(0.060)	(0.039)	(0.043)	(0.037)
Coors Light	0.569	−4.598	0.407	0.452	0.482
	(0.085)	(0.115)	(0.058)	(0.075)	(0.061)
Old Milwaukee Light . .	1.233	0.956	−6.097	0.841	0.565
	(0.121)	(0.132)	(0.140)	(0.112)	(0.087)
Lite	0.509	0.737	0.587	−5.039	0.577
	(0.095)	(0.122)	(0.079)	(0.141)	(0.083)
Molson Light	0.683	1.213	0.611	0.893	−5.841
	(0.124)	(0.149)	(0.093)	(0.125)	(0.148)

Hausman et al. Results

- Results: relatively high own- and cross-price elasticities (see tables).
- Counterfactuals: price changes post-merger
  - Markup for Miller Light goes up from 19.9% to 23.2% once other owned products are accounted for.
  - Note that Hausman et al. do not compute the new equilibrium, just the best response of merged firm to old competitor prices. (does this matter?)

Hausman (1997)

- Main issue: What is the value of a new good and how to incorporate it into the CPI?
- Basic idea is the use of the "virtual price", p^*. Then, we can calculate $e(p_{-n}, p_n^*, u)/e(p, u)$ which is the exact cost of living index. The value for consumers would be the area under the demand curve.
- Application is to the introduction of Apple-Cinnamon Cheerios.
- Data: weekly cereal data (p and q); 7 cities, 137 weeks.
- Demand: as above. Three segments (adults, kids, family) with 7, 4, and 9 brands each (so how much do we save over a full demand system?).

Hausman (1997) Estimates

Table 5.5 **Conditional Elasticities for Family Segment of RTE Cereal**

	Cheerios	Honey-Nut Cheerios	Apple-Cinnamon Cheerios	Corn Flakes	Kellogg's Raisin Bran	Rice Krispies	Frosted Mini-Wheats	Frosted Wheat Squares	Post Raisin Bran
Cheerios	−1.73851	0.16166	0.07110	0.19818	0.15355	0.09268	0.18649	0.02593	0.02716
	(0.04635)	(0.02520)	(0.01759)	(0.02776)	(0.02789)	(0.03008)	(0.02309)	(0.00628)	(0.02951)
Honey-Nut Cheerios	0.21637	−1.83838	0.14169	0.00390	0.18550	0.21253	0.04330	0.04414	0.09425
	(0.03686)	(0.05397)	(0.02562)	(0.03613)	(0.03750)	(0.03263)	(0.03967)	(0.01197)	(0.03863)
Apple-Cinnamon Cheerios	0.23945	0.34899	−2.11677	0.10597	0.23973	0.19848	0.01366	−0.02100	0.12936
	(0.06477)	(0.06330)	(0.07406)	(0.06451)	(0.06614)	(0.05520)	(0.06708)	(0.01993)	(0.06930)
Corn Flakes	0.23185	−0.03254	0.02883	−1.99465	0.23222	0.16056	0.07898	0.02246	0.13165
	(0.04859)	(0.04108)	(0.02952)	(0.06003)	(0.04533)	(0.03831)	(0.04307)	(0.01186)	(0.04444)
Kellogg's Raisin Bran	0.23744	0.21291	0.11121	0.28729	−1.94608	−0.08546	0.33045	0.06454	−0.10626
	(0.04839)	(0.04354)	(0.03084)	(0.04597)	(0.07233)	(0.04318)	(0.04474)	(0.01248)	(0.05031)
Rice Krispies	0.06656	0.19055	0.06997	0.16068	−0.12272	−2.00148	0.10508	−0.00909	0.34211
	(0.05873)	(0.04259)	(0.02824)	(0.04122)	(0.04759)	(0.06512)	(0.03707)	(0.00990)	(0.04614)
Frosted Mini-Wheats	0.43609	0.07708	0.00939	0.16386	0.48235	0.20255	−2.46950	0.14003	0.09692
	(0.05608)	(0.06460)	(0.04498)	(0.06371)	(0.06381)	(0.04921)	(0.11340)	(0.02669)	(0.06562)
Frosted Wheat Squares	0.37740	0.45906	−0.08636	0.26062	0.58179	−0.03396	0.86314	−3.16485	−0.09011
	(0.09617)	(0.12191)	(0.08357)	(0.11035)	(0.11175)	(0.08260)	(0.16566)	(0.13832)	(0.11552)
Post Raisin Bran	−0.10461	0.11474	0.08315	0.23661	−0.35988	0.73072	0.07025	−0.03721	−2.51416
	(0.12414)	(0.10689)	(0.07742)	(0.11177)	(0.12199)	(0.11060)	(0.10844)	(0.03036)	(0.15731)
Mean shares	0.21617	0.15026	0.06193	0.14243	0.13117	0.13539	0.09067	0.01475	0.05722

Note: Numbers in parentheses are asymptotic standard errors.

Hausman (1997), cont.

- Results: \$78M a year (30c per person), and CPI for cereals is over-estimated by roughly 25%.
- Caveats:
  - Functional form: can find a lower bound under convexity assumption (virtual price is about 35% lower).
  - Hausman also estimates a CES demand function, and finds the value three times higher (why?).
  - Instruments. Can we sign the potential bias?
  - Other prices are assumed fixed, but what if there are chosen strategically: CPI bias goes down to 20%, but this still does not take into account the reaction of other firms (in which direction will it go?).
- Bresnahan (1997): should we expect welfare effects of a new good in a segment to be large?