Static Models of Demand for Differentiated Products

Aviv Nevo

University of Pennsylvania and NBER

Fall 2018

Why Do We Care About Demand (in an IO course)?

- Allows us to "reverse engineer" firms' optimal decisions in order to
  - obtain marginal costs
  - test models of pricing
- Compute firm strategy that depends on consumer behavior
  - price discrimination
  - advertising and promotional activity
- Simulate counterfactuals
  - likely effect of mergers
  - demand for new products
- Consumer welfare

Background

The most straight-forward approach to specifying demand:

num products

$$q = D(p, r, \varepsilon)$$

q is a $J \times 1$ vector of quantities

p is a $J \times 1$ vector of prices

r is a vector of exogenous variables

ε is a $J \times 1$ vector of random shocks.

Early work focused on how 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; Barten 1966)

Translog model (Christensen, Jorgenson, and Lau, 1975)

Almost Ideal Demand System (Deaton and Muellbauer, 1980)

Issues

Issues for many cases considered in IO:

"The too many parameters problem". As the number of options, J, becomes large there are too many parameters to estimate

e.g., $D(p, r, \varepsilon) = Ap + \varepsilon$, where A is $J \times J$ matrix of parameters, implies J^2 parameters to be estimated.

with a more flexible functional form, the problem is even greater.

Issues

Heterogeneity. for some applications we would like to explicitly model and estimate the distribution of heterogeneity.

- above approach, generally, does not let us do this
- mere presence of heterogeneity doesn't invalidate aggregate demand
  - preferences are of the Gorman form (Gorman, 1959)
  - care in imposing the restrictions of economic theory

Issues

Explicit modeling of specific consumer behavior. easier when we start with an explicit model of consumer behavior and aggregate to market level aggregate demand.

New Goods. This demand system does not easily allow us to predict the demand for new goods.

Estimation. need to include, and instrument for, highly colinear prices.

"Solutions"

- Can generally be split into
  - product vs characteristics space
  - implicit vs explicit modeling of heterogeneity
- We will discuss:
  - aggregation across products
  - symmetry
  - weak separability and multi stage budgeting
  - models in characteristics space and discrete choice

Aggregation Across Products

- Aggregate individual products into aggregate commodities
- Can allow for flexible, even non-parametric, functional forms
- But for many IO problems aggregating to the level of the industry misses the point
- Real question is not whether to aggregate, but to what level and whether this aggregation solves the dimensionality problem
- The answer to how much to aggregate depends:
  - what we are interested in
  - correlation of prices and the substitution between the products we are aggregating over
    - how to compute an average price/price index
    - easy if aggregating over products with highly correlated prices
    - in general, need to know the substitution between the products

- Strong assump on demand or util

Symmetry Across Products

- Used mostly in the trade and applied theory literature
- Easy to work with analytically, but cannot fit many patterns in micro data
- A leading example: constant elasticity of substitution (CES) demand model. Let the utility from consumption of the J products be given by

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

where ρ is a constant parameter. The demand of the representative consumer:

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

where I is the income of the representative consumer.

A single parameter to estimate, regardless of the number of products

We solved the dimensionality problem by imposing symmetry between the different products:

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

- changing price of j, same effect globally
- Empirically CES is really strong assump.

Separability and Multi-Stage Budgeting

- Basic idea: solve the dimensionality problem by dividing the products into smaller groups and allowing for a flexible functional form within each group.
- Multi-stage budgeting: write the consumer's problem as a sequence of separate but related decision problems:
  - allocate expenditure to broad groups of products
  - allocate this expenditure to sub-groups of products, etc.
  - 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).
- Various conditions guarantee that the solution to this multi-stage process will equal the solution to the original consumer problem
- One condition is to assume weak separability of preferences.

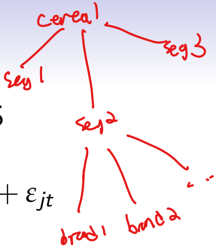
Weak Separability

- Let $\widetilde{q}_{1,} \widetilde{q}_2, ..., \widetilde{q}_G$ be G subvectors of the vector $\widetilde{q} = (q_1, q_2, ..., q_J)$ such that each product is only in one group. Then the utility is weakly separable if

$$U(\widetilde{q}) = f\left(v_1(\widetilde{q}_1), v_2\left(\widetilde{q}_2\right), ... v_G\left(\widetilde{q}_G\right)\right)$$

  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 process
- In order to justify that the higher stages of the decision process further assumptions are needed.
  - For example, indirect utility functions for each segment are of the Generalized Gorman Polar Form, and that the overall utility is separable additive in the sub-utilities

Application to DP



- Lowest level (demand for brand w segment): AIDS

$$s_{jt} = \alpha_j + \beta_j \ln(y_{gt}/\pi_{gt}) + \sum_{k=1}^{J_g} \gamma_{jk} \ln(p_{kt}) + \varepsilon_{jt}$$

where,

- s_{jt} dollar sales share of product j out of total segment expenditure
- y_{gt} overall per capita segment expenditure
- π_{gt} segment level price index
- p_{kt} price of product k in market t.

π_{gt} (segment price index) is either Stone logarithmic price index

$$\pi_{gt} = \sum_{k=1}^{J_g} s_{kt} \ln(p_{kt})$$

or

$$\pi_{gt} = \alpha_0 + \sum_{k=1}^{J_g} \alpha_k p_k + \frac{1}{2} \sum_{j=1}^{J_g} \sum_{k=1}^{J_g} \gamma_{kj} \ln(p_k) \ln(p_j).$$

"Demand for SUVs"

Multi-level Demand Model

- Middle level (demand for segments)

$$\ln(q_{gt}) = \alpha_g + \beta_g \ln(Y_{Rt}) + \sum_{k=1}^{G} \delta_k \ln(\pi_{kt}) + \varepsilon_{gt}$$

where

- q_{gt} quantity sold of products in the segment g in market t
- Y_{Rt} total category (e.g., cereal) expenditure
- π_{kt} segment price indices

Multi-level Demand Model

"demand for cars"

- Top level (demand for category)

$$\ln(Q_t) = \beta_0 + \beta_1 \ln(I_t) + \beta_2 \ln \pi_t + Z_t\delta + \varepsilon_t$$

where

- Q_t overall consumption of the category in market t
- I_t real income
- π_t price index for the category
- Z_t demand shifters

Models in Characteristics Space

Complaints: "How do you know segments are structured like this?"

- Up to now the model we focused on were in product space.
- We will turn to models based in characteristics space
- The utility function is a function of the attributes of the product (Gorman, 1956/1980, Lancester, 1966)
- Usually implemented as discrete choice, but does not have to be (active area of research)
- Connections to Hedonics

~Price index tools? Doesn't matter that much

Models in Characteristics Space

Indirect utility function of consumer i, from product j in market t:

$$U(x_{jt}, \xi_{jt}, I_i - p_{jt}, \tau_i; \theta),$$

where

x_{jt} – $1 \times K$ vector of observed product characteristics
ξ_{jt} – unobserved (by us) product characteristic
τ_i – individual characteristics
I_i – income

- ξ_{jt} will play an important role
    - realistic (inability of observed characteristics to capture the essence of the product(
    - will act as a "residual" (why don't predicted shares fit exactly – overfitting)
    - potentially implies endogeneity

Normalizations

- For what follows utility is cardinal, and therfore is invariant to affine transformation
- This means that we typically have 2 normalizations
- In what follows (will make more sense below)
    - normalize the variance of one component to one
    - normalize the utility from the outside good to zero

Linear RC (Mixed) Logit Model

A common model is the linear Mixed Logit model

$$u_{ijt} = x_{jt}\beta_i + \alpha_i(I_i - p_{jt}) + \xi_{jt} + \varepsilon_{ijt}$$

where ε_{ijt} is a stochastic term with $Var(\varepsilon_{ijt}) = 1$
first norm

- 2 views of ε_{ijt}:
  - utility is deterministic, but the choice process itself is probabilistic (Tversky, 1972)
  - utility is deterministic, but ε_{ijt} captures the researcher's inability to formulate individual behavior precisely
- Interplay between ξ_{jt} and ε_{ijt}: in a way all the ξ_{jt} is doing is changing the mean of ε_{ijt}, by j and t.

Heterogeneity

can put different weights on diff char and some put more emphasis on price.

Consumer-level taste parameters are modeled as

$$\alpha_i = \alpha + \sum_{r=1}^{d} \pi_{1r} D_{ir} + \sigma_1 v_{i1},$$

$$\beta_{ik} = \beta_k + \sum_{r=1}^{d} \pi_{(k+1)r} D_{ir} + \sigma_{k+1} v_{i(k+1)} \quad \text{for } k = 1, ..., K$$

where

$D_i = (D_{i1}, ..., D_{id})'$ is a $d \times 1$ vector of observed demographic variables

$v_i = (v_{i1}, ..., v_{i(K+1)})'$ is a vector of $K+1$ unobserved consumer attributes

Denote Π is a $(K+1) \times d$ matrix of parameters and $\sigma = (\sigma_1, ..., \sigma_{K+1})$ is a vector of parameters.

Outside Option

The indirect utility from this outside option is

$$u_{i0t} = \alpha_i I_i + \varepsilon_{i0t}$$

We normalize its mean utility to zero

second norm

The outside option will allow for substitution outside the market; important in many IO applications

- Let $\theta = (\overset{\text{MU from income}}{\alpha}, \overset{\text{char}}{\beta}, \overset{\text{Demog. int het.}}{\Pi}, \overset{\text{unobs het.}}{\sigma})$ denote the parameters of the model.
    - $\theta_1 = (\alpha, \beta)$ and $\theta_2 = (\Pi, \sigma)$
- It will be convenient to write

$$u_{ijt} = \underbrace{\delta_{jt}(x_t, p_t, \xi_t; \alpha, \beta)}_{\text{mean utility (across ind)}} + \mu_{ijt}(\underbrace{x_t, p_t, D_i; \Pi, \sigma}_{\text{het params}}) + \varepsilon_{ijt}$$

where $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ and

$$\begin{aligned} \mu_{ijt} \quad = \quad & -\left(\sum_{r=1}^{d} \pi_{1r} D_{ir} + \sigma_1 v_{i1} \right) p_{jt} + \\ & \sum_{k=1}^{K} \left(\sum_{r=1}^{d} \pi_{(k+1)r} D_{ir} + \sigma_{k+1} v_{i(k+1)} \right) x_{jt}^{k} \end{aligned}$$

$$u_{ijt} \equiv \delta(x_{jt}, p_{jt}, {}_{jt}; \theta_1) + \mu(x_{jt}, p_{jt}, D_i, \nu_i; \theta_2) + \varepsilon_{ijt}$$

Note:
(1) the mean utility will play a key role in what follows
(2) the interplay between μ_{ijt} and ε_{ijt}
(3) the "linear" and "non-linear" parameters
(4) definition of a market

Choice Probabilities and Market Shares

- Assume consumers purchase one unit of the good, which gives the highest utility.
- The probability that type (D_i, v_i) chooses option j is

$$s_{ijt} = s_{ijt}(x_t, \delta_t, p_t, D_i, v_i; \theta) = \int 1[u_{ijt} \geq u_{ikt} \ \forall l \mid x_t, \delta_t, p_t, D_i, v_i; \theta]$$

- We get market shares by integrating this probability over & integrating over ε's consumer attributes

$$s_{jt} = s_{jt}(x_t, \delta_t, p_t; \theta) = \int s_{ijt}(x_t, \delta_t, p_t, D_i, v_i; \theta) dF_D(D) dF_v(v)$$

- With consumer level data we will integrate only over v_i
- With assumptions on the distribution of the individual attributes can compute this integral

Logit

- The simplest assumptions we can make are
  1. $\Pi = 0$ and $\sigma = 0$, which implies $\beta_i = \beta$ and $\alpha_i = \alpha$
  2. ε_{ijt} are iid
  3. ε_{ijt} are distributed according to a Type I extreme value distribution.
- These imply

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

I is factored out because we only care about relative utility.

Price Elasticities

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

- 2 Problems
- Own price elasticities: market shares are small, so $\alpha(1 - s_{jt})$ is nearly constant and therefore the own-price elasticities are proportional to price.
    - driven mostly by lack of heterogeneity
- Cross-price elasticities: cross price elasticity wrt a change in the price of product k is that same for all products such that $j \neq k$.
    - driven by lack of heterogeneity and iid

Relaxing the iid Assumption

- Nested Logit:
  - $\Pi = 0$ and $\sigma = 0$,
  - divide the products into mutually exclusive nests, $g = 1, ..., G$.
  - let $\varepsilon_{ijt} = \lambda \varepsilon_{ig(j)t} + \varepsilon^1_{ijt}$, where ε^1_{ijt} is an iid extreme value shock, $\varepsilon_{ig(j)t}$ is a shock common to all options in segment g, and λ is a parameter that captures the relative importance of the two.
  - a particular distribution for $\varepsilon_{ig(j)t}$ gives the Nested Logit model.
  - if $\lambda = 0$ we get the Logit model.
- The Nested Logit model is a private case of the more general Generalized Extreme Value model which imposes correlation among the options through correlation in ε_{ijt}.

Introducing Heterogeneity

- Generate correlation through μ_{ijt} by allowing heterogeneity in tastes for the product attributes to drive correlation.
- For example, if "luxury" is an attribute of a car, then a consumer who likes one luxury car is more likely, then the average consumer to like another luxury car
- Nested Logit can be seen as a private case

Price Elasticities

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) dP_D(D) dP_v(v) & \text{if } j = k \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dP_D(D) dP_v(v) & \text{otherwise} \end{cases}$$