Discrete Choice Model

- Utility of consumer i for product j

$$u_{ij} = U(x_j, p_j, v_i)$$

  - x_j is a vector of product characteristics
  - p_j is the price of the product and
  - v_i is a vector of consumer characteristics

- Horizontal Model Example: Hotelling with quadratic costs

$$u_{ij} = \overline{u} - p_j - (x_j - v_i)^2$$

  - x_j is the location of the product along the line
  - v_i is the location of the consumer

Random Utility Model (RUM)(McFadden)

- U_{ij}^* usually specified as a sum of two parts

$$U_{ij}^*(x_j, p_j, p_z, y_i) = V_{ij}(x_j, p_j, p_z, y_i) + \varepsilon_{ij}$$

  - ε_{ij} i.i.d. across products and consumers; represents consumer tastes (observed by consumer but not by the researcher)
- What does it mean for tastes to be represented by product and consumer specific random terms?
  - product chosen is random from the researchers point of view
  - McFadden won the Nobel Prize for this in 2000
  - Assumptions about distribution of the ε_{ij}'s determines choice probabilities
- The probability that consumer i buys product j is

$$D_{ij}(p_1, ...p_j, p_z, y_i) = \Pr ob\left\{\varepsilon_{i0}, ..., \varepsilon_{ij} : U_{ij}^* > U_{ik}^*, \text{ for } j \neq k\right\}$$

Independence of Irrelevant Alternatives (IIA)

- ratio of choice prob (odds ratio) does not depend on the number of alternatives available

$$\frac{s_{ij}}{s_{in'}} = \frac{\exp(V_{ij})}{\exp(V_{in'})}$$

- Red bus/blue bus problem: Walk or take red bus
  - If consumer walks half the time then $s_{iW} = s_{iRB} = 0.5$
  - odds ratio walk/RB=1
- Introduce a red bus
  - odds ratio between walk/BB is 1
- But buses are perfect substitutes
  - new choice prob should be $s_{iW} = 0.5; s_{iRB} = s_{iBB} = 0.25$
  - new odds ratio should be walk/RB=2
- IIA is especially troubling if want to predict penetration of new products

Independence of Irrelevant Alternatives (IIA)

- ratio of choice prob (odds ratio) does not depend on the number of alternatives available

$$\frac{s_{ij}}{s_{in'}} = \frac{\exp(V_{ij})}{\exp(V_{in'})}$$

- Red bus/blue bus problem: Walk or take red bus
  - If consumer walks half the time then $s_{iW} = s_{iRB} = 0.5$
  - odds ratio walk/RB=1
- Introduce a red bus
  - odds ratio between walk/BB is 1
- But buses are perfect substitutes
  - new choice prob should be $s_{iW} = 0.5; s_{iRB} = s_{iBB} = 0.25$
  - new odds ratio should be walk/RB=2
- IIA is especially troubling if want to predict penetration of new products

Counter-intuitive substitution patterns:

- Not only from the distributional logit assumption
- Due to assumption that the only variance in consumer tastes comes through the i.i.d. product-specific terms ε_{ij}
- Since i.i.d., there is no source of correlation in consumer tastes across similar products
- Changes to allow for more intuitive substitution patterns
  - Generalized EV models (GEV, Nested logit)
  - Mixtures of logits (K types of logit parameters)
  - Product differentiation model (Bresnahan, Stern, Trajtenberg 1997)
  - Random Coefficients Model of Demand (Berry, Levinsohn, and Pakes)

Nested Logit Model

- Within-group correlation parameter is σ_g
- Across nests, parameter σ (within (0,1)) describes correlation between nests

$$u_{ij} = x_j\beta - \alpha p_j + \sigma_g v_{ig} + \varepsilon_{ij}$$

- Define the inclusive value of nest g as:

$$s_{ig} = \sum_{j \in g} \exp\left(\frac{u_{ij}}{1-\sigma}\right)$$

- McFadden (1978) showed nested structure is consistent with RUM maximization iff the coefficients of the inclusive value lie within the unit interval
- More complicated forms of cross-product correlation in tastes do not lead to closed form expressions for shares (like Nested Logit does)
    - need to compute a high dimensional integral and this is tough
    - simulation methods help here

Inversion Example: Berry Logit

- Simple example of the inversion step: MNL shares

$$\widehat{s}_{jt}(\delta) = \frac{\exp(\delta_{jt})}{1 + \sum_{k=1}^{J} \exp(\delta_{kt})}$$

$$\log \widehat{s}_{jt} = \delta_{jt} - \log \left(1 + \sum_{k=1}^{J} \exp(\delta_{kt}) \right)$$

- but notice for outside good

$$\log \widehat{s}_{0t} = 0 - \log \left(1 + \sum_{k=1}^{J} \exp(\delta_{kt}) \right)$$

- so $\delta_{jt} = \log \widehat{s}_{jt} - \log \widehat{s}_{0t}$

- This implies

$$(\log S_{jt} - \log S_{0t}) = \delta_{jt} = X_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

  - where S_{0t} is the share of the outside good
- Can be estimated by OLS
  - dependent variable $\log S_{jt} - \log S_{0t}$
  - covariates X_{jt}, p_{jt}, and error term: ξ_{jt}
- When there is not a closed form solution for the market share then solve for ξ structural error and construct moment condition
- Restrict the model predictions for product j's market share to match the observed market shares

$$S_t^{obs} - s_t(\delta, \theta) = 0$$

- then solve for the demand side unobservable

$$\xi_{jt} = \delta_{jt}(S, \theta) - X_{jt}'\beta - \alpha p_{jt}$$

- This implies

$$(\log S_{jt} - \log S_{0t}) = \delta_{jt} = X_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

  - where S_{0t} is the share of the outside good
- Can be estimated by OLS
  - dependent variable $\log S_{jt} - \log S_{0t}$
  - covariates X_{jt}, p_{jt}, and error term: ξ_{jt}
- When there is not a closed form solution for the market share then solve for ξ structural error and construct moment condition
- Restrict the model predictions for product j's market share to match the observed market shares

$$S_t^{obs} - s_t(\delta, \theta) = 0$$

- then solve for the demand side unobservable

$$\xi_{jt} = \delta_{jt}(S, \theta) - X'_{jt}\beta - \alpha p_{jt}$$

What are appropriate instruments?

- IV for j should be correlated with p_j but not with structural error ξ_j
- Usual demand case: cost shifters
  - but we have cross-sectional (across products) data, so we require IV to vary across products within a market
- Example: cars, one natural cost shifter are wages in Michigan
- Here doesn't work because its the same across all products
  - if ran 2SLS with wages in Michigan as IV, first stage regression of price on wage would yield the same predicted price for all products

- One commonly used specification is the logit model with random (normal) coefficients

$$U_{ij} = X_j \beta_i - \alpha p_j + \xi_j + \varepsilon_{ij}$$

  - The K random coefficients (one for each product characteristic) are

$$\beta_{ik} = \beta_k + \sigma_k v_{ik}$$
$$v_{ik} \sim N(0,1), iid$$

  - it is useful to decompose utility into two parts

$$\mu_{ij} = \Sigma_k \sigma_k X_{jk} v_{ik}$$
$$\delta_j = X_j \beta_k - \alpha p_j + \xi_j$$

  - So we can rewrite indirect utility as

$$u_{ij} = \delta_j + \mu_{ij} + \varepsilon_{ij}$$

Intuition of the Estimation Algorithm

- The model is one of individual behavior, yet only aggregate data is observed.
- We can still estimate the parameters that govern the distribution of individuals
  - compute predicted individual behavior and aggregate over individuals, for a given value of the parameters,
  - obtain predicted market shares
- We then choose the values of the parameters that minimize the distance between these predicted shares and the actual observed shares
- The metric under which this distance is minimized is not the straightforward sum of least squares
- rather it is the metric defined by the instrumental variables and the GMM objective function
- It is this last step that somewhat complicates the estimation procedure

Intuition of the Estimation Algorithm

- The model is one of individual behavior, yet only aggregate data is observed.
- We can still estimate the parameters that govern the distribution of individuals
  - compute predicted individual behavior and aggregate over individuals, for a given value of the parameters,
  - obtain predicted market shares
- We then choose the values of the parameters that minimize the distance between these predicted shares and the actual observed shares
- The metric under which this distance is minimized is not the straightforward sum of least squares
- rather it is the metric defined by the instrumental variables and the GMM objective function
- It is this last step that somewhat complicates the estimation procedure

Calculating the market share via simulation:

- More detail on step 1:
- Condition on v_i, y_i – this is a logit and get closed form
- Take draws on v_i, y_i and average over the implied logit shares:

$$\sum_{i=1}^{ns} \frac{\exp(\overline{\mu}_{ij} + \delta_j)}{\Sigma_k \exp(\overline{\mu}_{ik} + \delta_k)}$$

- BLP then provide an algorithm (a contraction mapping) that solves for δ given the parameters and a set of simulation draws

Having discussed some methods of deriving demand, we now turn to the supply side and a consideration of equilibrium by considering in turn:

- Estimation of supply side parameters
- Incorporating multi-product firms
- Simultaneously estimating supply and demand
- How to estimate degree of market power or presence of collusion

Supply Side

- Simplest models of product differentiation involve 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.
- 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

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

- We can rewrite as $p_j = mc_j + b_j(p)$

- where the price-cost markup is

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

- Assume that marginal cost is

$$mc_j = w_j\eta + \lambda q_j + \omega_j$$

  - where w_j might consist of X and input prices and q is output
  - ω_j is a supply shock unobserved to the econometrician
- Combining, the FOC is then

$$p_j = w_j\eta + \lambda q_j + b_j(p) + \omega_j$$

- If demand parameters are known then the markup is known and can estimate by IV methods (eg 2SLS) where IV are demand-side variables
- Alternatively mc and demand can be estimated together

Multi-Product Firms

- Non-cooperative oligopolistic Bertrand competition
- Firm f produces a subset $j \in \mathcal{J}_f$ of the products: Profits

$$\sum_{j \in \mathcal{J}_f} (p_j - mc_j) \mathcal{M} s_j(p, X, \xi; \theta)$$

  - where $\mathcal{M}$ is market size
  - s_j is the simulated aggregate market share
- Marginal costs

$$mc_j = w_j' \eta + \omega_j$$

- Any product must have prices that satisfy

$$s_j(p, a) + \sum_{r \in \mathcal{J}_f} (p_r - mc_r) \frac{\partial s_r(p, a)}{\partial p_j} = 0$$

- Given demand can solve for marginal costs and for ω_j

- In vector form, the J FOC are

$$s - \Omega(p - mc) = 0$$

- Notice this implies a markup equation $p - mc = \Omega^{-1}s$
- Ω is called the ownership matrix (of dimension JxJ)
  - Each element takes on the value of $\partial s_r(p, a)/\partial p_j$ for every product that the firm owns
- To estimate the FOC think of estimating the equation

$$mc_j = p_j - b_j(p, x, \xi; \theta) = w_j'\eta + \omega_j$$

- Just as in estimating demand, estimates of the parameters η can be obtained from orthogonality conditions between ω and appropriate instruments

- In vector form, the J FOC are

$$s - \Omega(p - mc) = 0$$

- Notice this implies a markup equation $p - mc = \Omega^{-1}s$
- Ω is called the ownership matrix (of dimension JxJ)
  - Each element takes on the value of $\partial s_r(p, a)/\partial p_j$ for every product that the firm owns
- To estimate the FOC think of estimating the equation

$$mc_j = p_j - b_j(p, x, \xi; \theta) = w_j'\eta + \omega_j$$

- Just as in estimating demand, estimates of the parameters η can be obtained from orthogonality conditions between ω and appropriate instruments

Estimation of Supply and Demand Side

- Demand side moment: Restrict the model predictions for product j's market share to match the observed market shares

$$S_t^{obs} - s_t(\delta, \theta) = 0$$

- then solve for the demand side unobservable

$$\xi_{jt} = \delta_{jt}(S, \theta) - x_j' \beta$$

- Cost side moment:
- Rearranging price FOC's yields

$$mc = p - \Omega^{-1} s$$

- combined with marginal costs yields cost side unobservable

$$\omega = \ln(p - \Omega^{-1} s) - w' \eta$$

Nevo: Measuring Market Power in the RTE Cereal Industry

- The ready-to-eat (RTE) cereal industry is characterized by high price-to-cost margins (PCM) and high concentrations
- Antitrust authorities accused firms of collusive pricing behavior
- Nevo tests whether this is the case by estimating the price-cost margin (PCM) and decomposing it into 3 sources:
    1. that due to product differentiation
    2. that due to multiproduct form pricing and
    3. that due to price collusion
- Overview of methodology:
    - use the BLP framework to estimate brand-level demand.
    - use demand estimates and different pricing rules to back out PCMs.
    - compare PCMs against crude measures of actual PCM to separate the different sources of the markup

Model and Data

- Indirect utility is

$$u_{ijt} = \alpha_i p_{jt} + X_j \beta_i + \xi_j + \Delta\xi_{jt} + \varepsilon_{ijt}$$

  - uses brand dummy variables (ξ_j) to capture the mean characteristics of RTE cereal
  - once brand dummy variables are included in the regression, the error term is the unobserved city-quarter specific deviation from the overall mean valuation of the brand :structural error is the change in ξ_j over time (denoted $\Delta\xi_{jt}$)
- Cannot use BLP Type Instruments
  - there is no variation in each brand's observed characteristics over time and across cities
  - only variation in IVs from characteristics is due to changes in choice set of available brands
  - proposes alternative IV to separate the exogenous variation in prices (due to differences in mc) and endogenous variation (due to differences in unobserved valuation)

IVs with brand dummies

- Exploit the panel structure of the data (similar to those used by Hausman (1996))
- The identifying assumption is that, controlling for brand specific means and demographics, city-specific valuations are independent across cities (but are allowed to be correlated within a city)
- Given this assumption, the prices of the brands in other cities are valid IV's.
  - prices of brand j in two cities correlated due to the common mc
  - but due to the independence assumption will be uncorrelated with market specific valuation.
  - One could potentially use prices in all other cities and all quarters as instruments
- Independence assumption may not hold (for instance, if there is a national demand shock related to health of cereal)

Identifying Collusive Behavior

- Recall the markup is given by

$$p - mc = \Omega^{-1} s$$

- With single product firms the price of each brand is set by a profit-maximizing firm that considers only the profits from that brand. In this case the ownership matrix will be diagonal
- With multi-product firms, firms set the prices of all their products jointly. In this case some off diagonals will be non-zero
- With collusion, firms act as one firm which owns all products (ie joint profit-maximization of all the brands). In this case the ownership matrix will have no zeros
- Nevo estimates parameters under different definitions of the ownership matrix

Identifying Collusive Behavior

- Recall the markup is given by

$$p - mc = \Omega^{-1} s$$

- With single product firms the price of each brand is set by a profit-maximizing firm that considers only the profits from that brand. In this case the ownership matrix will be diagonal
- With multi-product firms, firms set the prices of all their products jointly. In this case some off diagonals will be non-zero
- With collusion, firms act as one firm which owns all products (ie joint profit-maximization of all the brands). In this case the ownership matrix will have no zeros
- Nevo estimates parameters under different definitions of the ownership matrix

Results

- Compares predicted PCM under all three situations to the observed PCM calculated using accounting data for costs
- Finds that the first two effects explain most of the observed price-cost margins
- Prices in the industry are consistent with noncollusive pricing behavior, despite the high price-cost margins.