STRUCTURE

Aim is the applied side of some of the techniques you have been learning from Prof. Pischke.

The empirical counterpart to core work you have been doing on strategy (Prof Sutton) and firm organization (Prof Garicano).

**Lectures:** I will post all the lectures on my website (under “teaching”) & on Moodle:

http://cep.lse.ac.uk/_new/staff/person.asp?id=1358

**Breaks:** I’ll take a 5 minute break around 11am
STRUCTURE OF THE LECTURES

- **Demand**
  - Demand estimation of single product (lecture 1)
  - Demand estimation of differentiated products (lecture 2, today)

- **Productivity - Technology & Management Practices**
  - Production Functions (lecture 3)
  - Management I - measurement & drivers (lecture 4)
  - Management II - effects on productivity (lecture 5)

- **Strategy topics**
  - Price Discrimination (lecture 6)
  - Entry (lecture 7)

- **Organization topics**
  - Complementarities (lecture 8)
  - Decentralization (lecture 9)
  - [Information (lecture 10)]
STRUCTURE OF THE CLASSES

- **Alexander Lembcke** ([A.C.Lembcke@lse.ac.uk](mailto:A.C.Lembcke@lse.ac.uk))
- Every week starting today. Problem sets will be on Moodle. Classes follow lectures with a lag
  - E.g. Demand (3 classes); Productivity (2 classes)
- These will be usually data analysis exercises
- **Readings**
  - Books. Econometrics. Use Stock and Watson as basic but Wooldridge better (advanced); Church & Ware (strategy/IO; lots of examples);
  - Papers: Starred readings best
    - E.g. Demand (Ackerberg et al part 1); Productivity (Ackerberg et al part 2), management (Bloom & Van Reenen, 2007), etc.
DIFFERENTIATED PRODUCTS
Demand model

So far: Estimating Models of Demand and Supply with Homogenous Products

Now: Differentiated Products

Plan

1. Theoretical Summary
2. Pioneering work by McFadden
3. Recent Applications: BLP and Nevo
4. Some Remarks on other Approaches
Review of Last Lecture on Demand

- Why estimate demand curves?
Review of Last Lecture on Demand

- **Why estimate demand curves?**
  - Business (Price discrimination; M&A; ads; etc.); Policy (taxes; inflation; etc.)

- **Basic Estimation issues**
  - Elasticity

- **Supply side**
  - Costs
    - Imperfect competition (mark-ups)

- **Instrumental variables** (price endogeneity)
  - Cost shifters
Overview of Today’s lecture

- Continuous demand model
- Basic Discrete Choice model
- Beyond simple logit (Relaxing IIA)
  - Nested Logit
  - Mixed Logit/BLP
- Aggregation
- Endogeneity
- Applications
Continuous demand model

- Empirical specifications are based on these theoretical elements
- Representative consumer model
  - Specify the demand for an individual product
- Example 1: AIDS (almost ideal demand system), Deaton and Muellbauer
- Budget share equation
  \[ w_i = \alpha_i + \sum \gamma_{ij} \log p_j + \beta_i \log(x/P) \]
  where \( x \) is total expenditures and \( P \) is the price index defined

- \( w_i \) is the expenditure share of individual product \( i \)
Continuous demand model

- Empirical specifications are based on these theoretical elements
- Representative consumer model
  Specify the demand for an individual product
- Example 1: AIDS (almost ideal demand system), Deaton and Muellbauer
- Budget share equation

\[ w_i = \alpha_i + \sum \gamma_{ij} \log p_j + \beta_i \log(x/P) \]

where \( x \) is total expenditures and \( P \) is the price index defined

Own price, \( \gamma_{ii} \)  
Cross- Prices of all other substitutes/complements, \( \gamma_{ij} \)

- \( w_i \) is the expenditure share of individual product i
Advantages of continuous demand

- Relatively easy to calculate the elasticity
  - Note: elasticity not simply the coefficient on price as dependent variable is a share, not log(quantity)
- Easy to impose economic restrictions on system of demand estimation (symmetry, homogeneity, etc.)
- Flexible – allows free pattern of substitution across brands/products
- Can add multiple levels (Gorman “two stage budgeting”)
- Do not need to define relative characteristics of each brand (cf. and discrete choice approach)
Limitations of continuous demand model

1. Representative consumer
2. Curse of Dimensionality
3. Price Endogeneity
4. New Goods
Limitations of continuous demand model: (1/4) Representative consumer assumption

- **Representative consumer assumption**
  - Obviously untrue
  - Demand for some goods depends not only on aggregate income but on distribution of income & other variables affecting preferences (e.g. age & education)
  - But substitution possibilities also differ across consumers
  - No corner solutions: everyone buys at least a little

- **Common “fix up”**
  - Add in controls for higher order “moments” of income and other variables (e.g. Variance of income)
  - But (a) ad hoc and (b) limited #s included because of “curse of dimensionality”/too many parameters (see below)
Limitations of continuous demand model: (2/4) “Curse of dimensionality”

- For every good we have to include the prices of both the good itself and (approx) every other good.
- Since the number of products (J) can be large in some applications (e.g. 100s of cars) the number of parameters > number of observations
- Example: we need to estimate K parameters where \( K = 2J + (J(J+1))/2 \), so if \( J=100 \) \( K=5,250 \) coefficients!!
- In practice we reduce the number by:
  - imposing some assumptions (e.g. symmetry and homogeneity)
  - Aggregating into larger groups (e.g. Foreign cars vs. domestic cars, luxury vs. mass market cars)
  - But still a lot and aggregating can become ad hoc
    - Marketing literature can help here, e.g. RTE cereal division
An Application of Continuous Demand (Hausman, 1997, on RTE cereals)

- Ready to Eat breakfast cereals 7 US cities; 137 weekly obs 1990-92
- 3 level demand system (TOP: all cereal; MIDDLE cereal segments; LOWER: brand level (up to 9 brands per segment)
- **Lowest** level $w = \text{Share in city } n \text{ of a cereal brand } i \text{ in total cereal segments}$
  - Adult (7 brands - Shredded Wheat, Special K, Grape Nuts...)
  - Child (4 - Frosted Flakes, Froot Loops, other)
  - Family (9 - Cheerios, Corn Flakes, Rice Crispies, ...)
  - Total spending on segment cereal, deflated by segment price Index
- **Middle** level is share of segment (e.g. Adult) in total cereal. Uses simple log-log model; expenditure on all cereals
- **Upper** is total cereal demand as a function of cereal price index and total consumer demand (estimated on a 16 year time series)
- **Problems:** (i) ad hoc; (ii) weak identification of all the cross price elasticities; (iii) too many brands? (“other” category)
Some results from Hausman (1997)

<table>
<thead>
<tr>
<th></th>
<th>Cheerios</th>
<th>Honey-Nut Cheerios</th>
<th>Apple-Cinnamon Cheerios</th>
<th>Corn Flakes</th>
<th>Kellogg’s Raisin Bran</th>
<th>Rice Krispies</th>
<th>Frosted Mini-Wheats</th>
<th>Frosted Wheat Squares</th>
<th>Post Raisin Bran</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheerios</td>
<td>-1.73851</td>
<td>0.16166</td>
<td>0.07110</td>
<td>0.19818</td>
<td>0.15355</td>
<td>0.09268</td>
<td>0.18649</td>
<td>0.02593</td>
<td>0.02716</td>
</tr>
<tr>
<td>(0.04635)</td>
<td>(0.02520)</td>
<td>(0.01759)</td>
<td>(0.02776)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honey-Nut Cheerios</td>
<td>0.21637</td>
<td>-1.83838</td>
<td>0.14169</td>
<td>0.09290</td>
<td>0.18550</td>
<td>0.21253</td>
<td>0.04330</td>
<td>0.04414</td>
<td>0.09425</td>
</tr>
<tr>
<td>(0.03686)</td>
<td>(0.05397)</td>
<td>(0.02562)</td>
<td>(0.03613)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple-Cinnamon Cheerios</td>
<td>0.23945</td>
<td>0.34899</td>
<td>-2.11677</td>
<td>0.10597</td>
<td>0.23973</td>
<td>0.19848</td>
<td>0.01366</td>
<td>-0.02100</td>
<td>0.12936</td>
</tr>
<tr>
<td>(0.06477)</td>
<td>(0.06330)</td>
<td>(0.07406)</td>
<td>(0.06451)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corn Flakes</td>
<td>0.23185</td>
<td>-0.03254</td>
<td>0.02883</td>
<td>-1.99465</td>
<td>0.23222</td>
<td>0.16056</td>
<td>0.07898</td>
<td>0.02246</td>
<td>0.13165</td>
</tr>
<tr>
<td>(0.04859)</td>
<td>(0.04108)</td>
<td>(0.02952)</td>
<td>(0.06003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kellogg’s Raisin Bran</td>
<td>0.23744</td>
<td>0.21291</td>
<td>0.11121</td>
<td>0.28729</td>
<td>-1.94608</td>
<td>-0.08546</td>
<td>0.33045</td>
<td>0.06454</td>
<td>-0.10626</td>
</tr>
<tr>
<td>(0.04839)</td>
<td>(0.04354)</td>
<td>(0.03084)</td>
<td>(0.04597)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice Krispies</td>
<td>0.06656</td>
<td>0.19055</td>
<td>0.06997</td>
<td>0.16068</td>
<td>-0.12272</td>
<td>-2.00148</td>
<td>0.10508</td>
<td>-0.09090</td>
<td>0.34211</td>
</tr>
<tr>
<td>(0.05873)</td>
<td>(0.04259)</td>
<td>(0.02824)</td>
<td>(0.04122)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frosted Mini-Wheats</td>
<td>0.43609</td>
<td>0.07708</td>
<td>0.09399</td>
<td>0.16386</td>
<td>0.48235</td>
<td>0.20255</td>
<td>-2.46950</td>
<td>0.14003</td>
<td>0.09692</td>
</tr>
<tr>
<td>(0.05608)</td>
<td>(0.06460)</td>
<td>(0.04498)</td>
<td>(0.06371)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frosted Wheat Squares</td>
<td>0.37740</td>
<td>0.45906</td>
<td>-0.06626</td>
<td>0.26062</td>
<td>0.58179</td>
<td>-0.03390</td>
<td>0.86314</td>
<td>-3.16485</td>
<td>-0.09011</td>
</tr>
<tr>
<td>(0.09617)</td>
<td>(0.12191)</td>
<td>(0.08357)</td>
<td>(0.11035)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post Raisin Bran</td>
<td>-0.10461</td>
<td>0.11474</td>
<td>0.08315</td>
<td>0.23661</td>
<td>-0.35988</td>
<td>0.73072</td>
<td>0.07025</td>
<td>-0.03721</td>
<td>-2.51416</td>
</tr>
<tr>
<td>(0.12414)</td>
<td>(0.10689)</td>
<td>(0.07742)</td>
<td>(0.11177)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean shares</td>
<td>0.21617</td>
<td>0.15026</td>
<td>0.06193</td>
<td>0.14243</td>
<td>0.13117</td>
<td>0.13359</td>
<td>0.09067</td>
<td>0.01475</td>
<td>0.05722</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are asymptotic standard errors.
Limitations of continuous demand model (3/4): Endogeneity of price

- We need an instrumental variable for price
- With multiple prices we need an instrument for every price.
- Ideally, shocks to the brand specific cost structure could be used, but these are not usually available. We may have some city-specific information on costs (e.g. wages) or aggregate costs, but these usually don’t vary between brands.
- Solution? Hausman, Leonard and Zona (1994) on beer “prices in other cities”
  - Idea is that brand-specific cost shocks will be reflected in higher prices across all cities
  - Problem is that there may also be unobserved demand shocks (e.g. Advertising or tastes) which may increase demand and price. So “other cities” no excludable (Bresnahan comment on Hausman, 1997)
Limitations of continuous demand model (4/4) New Goods

- Cannot easily analyze introduction of new goods
- To analyze the demand for a new good we really need to know parameters associated with this good (e.g. the own and cross price elasticities)
Overview

- Continuous demand model
- **Basic Discrete Choice model**
  - Random utility model
  - McFadden & multinomial logit
  - Applications
  - Disadvantages (IIA)
- Beyond simple logit (Relaxing IIA)
- Aggregation
- Endogeneity
- Applications
Random Utility model

- Conditional utility (for each good $j$ consumer $i$):

$$U_{ij} = \alpha C_i + x_{ij} \beta + \varepsilon_{ij}$$

- It allows for horizontal differentiation in consumer taste

- Consumer maximizes utility subject to budget constraint

$$\max_j U_{ij} \text{ subject to } p_j + C_i \leq y_i$$

$$U_{ij} = \alpha (y_i - p_j) + x_{ij} \beta + \varepsilon_{ij}$$
Discrete choice models describe decision among alternatives.

The choice set needs to exhibit three characteristics:
- the alternatives must be mutually exclusive
- the choice set must be exhaustive
- the number of alternative must be finite
Discrete choice models

- Only the difference in utility matters

\[ x_{ij} \beta - \alpha p_j + \varepsilon_{ij} + c > x_{ik} \beta - \alpha p_k + \varepsilon_{ik} + c \]

- Then we need to normalize the utility of a good. Usually the utility of the outside option is set equal to 0

- the scale is arbitrary

\[ \frac{x_{ij} \beta - \alpha p_j + \varepsilon_{ij}}{c} > \frac{x_{ik} \beta - \alpha p_k + \varepsilon_{ik}}{c} \]

- We normalize standard deviation of the error term
Discrete choice models

- McFadden won Nobel price for pioneering work on discrete choice models
- Two Key Assumptions
  1. Each consumer buys at most one unit
  2. Utility is determined by the characteristics of the product
Discrete choice models

Example 2: Multinomial Logit product choice

(i) Consumer $i$ chooses one from 1, ..., $m_i$ products

(ii) Utility of product $k$

\[ U_{ik} = x'_{ik} \beta - \alpha p_k + \varepsilon_{ik} \]

\[ = \delta_{ik} + \varepsilon_{ik} \]

where the variables have the following meaning:

- $x_{ik}$ are explanatory variables including product characteristics,
- $p_k$ is the price of product $k$,
- $\delta_{ik}$ is the 'mean utility' and
- $\varepsilon_{ik}$ are iid log Weibull distributed

\[ \text{extreme value} \rightarrow \exp(-\exp(-\varepsilon_{ik})) \]
Discrete choice models

- McFadden shows the following
- Result: The probability of choice \( k \) by consumer \( i \) equals

\[
P_{ik} = \frac{\exp(\delta_{ik})}{\sum_{j=0}^{m_i} \exp(\delta_{ij})}
\]

and is derived from utility maximization if and only if \( \varepsilon_{ik} \) are iid log Weibull distributed
Empirical implication

- Aggregating purchase probabilities these will map into (potential) market shares, $s_{ik}$, ($L =$ market size)

\[
\ln s_{ik} = \ln \left( \frac{q_{ik}}{L} \right) = -\alpha p_k + x_{ik}' \beta + \varepsilon_{ik} = \delta_{ik} + \varepsilon_{ik}
\]

- Notice that this equation has only one price and a vector of characteristics of (own good) on right hand side

- Goods have all been mapped into characteristic space (Lancaster)

- Under continuous demand we had all prices and a system of $K$ equations!

- But this simplicity comes at a price (no free lunch in econometrics)
Discrete choice models

- **Power**
  - Simplicity
  - It can represent systematic taste variation (linked to observable characteristic like with micro data), but not random variation in taste

- **Limitations of the logit model:**
  - It cannot deal with unobservable factors that are correlated over time
  - Independence of irrelevant alternatives (IIA)
Independence of irrelevant alternatives

- The probability ratio between two alternatives is independent from the presence of other alternatives

\[
\frac{P_{ij}}{P_{ik}} = \frac{\sum_{k=0}^{m_i} \exp(\delta_{ik})}{\sum_{k=0}^{m_i} \exp(\delta_{ik})} = \exp(\delta_{ij} - \delta_{ik})
\]

- Note that no other prices or characteristics from any other product except j and k (unlike continuous choice)
Independence of irrelevant alternatives

- The cross elasticity is the same for all alternatives. A change in an attribute (z or x) of alternative $k$ changes the probabilities for all the alternatives by the same percentage.

\[ E_{jz_{ik}} = -\beta z_{ik} P_{r_{ik}} \]

- In other words, the cross price elasticities all depend on market shares ($P_{r_{ik}}$) rather than being estimated from response to price variation in data. Very unsatisfactory:
  - Example: in merger analysis only market shares are relevant. But some products will be much closer substitutes than others even if small market shares
Example (McFadden, 1984): Red bus/blue bus

- Suppose there are just two travel choices:
  - taking a car (“c”) or a blue bus (“bb”).
- Assume that $Pr_c = 1/2$, $Pr_{bb} = 1/2$ so that $Pr_c / Pr_{bb} = 1$.
- If a red bus (“rb”) is introduced because the colour of the bus will not affect the probability of taking it with respect to the blue bus we must have that $P_{bb} = Pr_{rb}$
- However, IIA implies that $Pr_c / Pr_{bb}$ will not change.
- Together the two equalities implies that $Pr_{bb} = Pr_{rb} = Pr_c = 1/3!$
  - Clearly unsatisfactory since buses must be closer substitutes with each other than cars for buses.
  - IIA driven by strict functional form assumptions of logit (extreme value/log Weibull distribution)
Testing for IIA:

- Estimate a model for a subset of alternatives.
- Under IIA, the probability ratio between the two given alternatives is not different whether or not the other alternatives are included.
- In practice you compare the estimates under the two alternatives and do a Hausman Test
- Hausman-McFadden (1984), Wooldridge p.502
What can we do to relax IIA?

- Nested Logit (a “hierarchical model”)
- Mixed Logit / Random coefficient models (BLP)
Overview

- Continuous demand model
- Basic Discrete Choice model
- **Beyond simple logit (Relaxing IIA)**
  - Nested Logit
  - Mixed Logit/BLP
- Aggregation
- Endogeneity
- Applications
Nested logit

- A nested logit is appropriate when the set of alternatives can be partitioned into (more homogeneous) subsets (nests).

- Example
  - Car/motorbike (private transportation)
  - Bus/Train (public transportation)
  - Walking (outside option)
Nested logit

- The model has the following properties:
  - For any two choices in the same nests IIA holds
  - For any two choices in different nests IIA does not hold
Nested logit

\[
Pr_{ij} = \frac{\exp(\delta_{ij}/\lambda_k) \left( \sum_{j \in B_k} \exp(\delta_{ij}/\lambda_k) \right)^{\lambda_k^{-1}}}{\sum_{l} \left( \sum_{j \in B_l} \exp(\delta_{ij}/\lambda_l) \right)^{\lambda_l}}
\]

- Notice that for $\lambda_k = 1$ we have the standard logit model. We need $\lambda_k \geq 0$ for the model to be consistent with utility maximization.
Examples of nested Logit

- Goldberg (1995) car demand, Household data

Classes: compacts, subcompacts, Intermediate, luxury, sports, pick-ups, Vans, other

Figure 1.—Automobile choice model.
Examples of nested Logit

- Ivaldi and Verboven (2005) on European trucks
- Volvo-Scania merger. First time discrete choice econometrics used in a major EU merger case
- Predicted price increases of 5% in EU & 10% in Scandinavia
Beyond Nested Logit: Berry, Levinsohn & Pakes (1995)

- Nested logit still puts strong conditions on demand elasticities. Within a nest/group the IIA assumption holds so elasticities will just be determined by market shares again.
- Deciding on which nests is somewhat arbitrary. Same issue as in continuous demand when deciding number of levels and groups.
- Ordering of nests matters (e.g. Goldberg: do people decide foreign/domestic first or van/car first?). Can test but gets more complicated and arbitrary.
Mixed Logit/Random Coefficients: BLP

- Allow for consumers to have different valuation of price (p) and characteristics (x)
  \[ U_{ij} = \alpha_i (y_i - p_j) + x_{ij} \beta_i + \varepsilon_{ij} \]
- The \( \alpha_i \) and \( \beta_i \) assumed to be random variables (e.g. multivariate normal)
- Allows much more flexible patterns of substitution and realistic demand elasticities
- But cannot express market shares analytically. Have to estimate by simulation rather than OLS/IV
Overview

- Continuous demand model
- Basic Discrete Choice model
- Beyond simple logit (Relaxing IIA)
  - Nested Logit
  - Mixed Logit/BLP
- Aggregation
- Endogeneity
- Applications
Aggregate data

- Conditional utility:

\[ U_j = x_j \beta - \alpha p_j + \epsilon_{ij} \]

- By estimating the model under the logit assumption we usually get implausible estimates (positive price coefficient)
Aggregate data (market shares rather than individual choices)

- Logit model implies:

\[ P_j = \frac{e^{x_j \beta - a p_j + \xi_j}}{\sum_j e^{x_j \beta - a p_j + \xi_j}} \]

- It is likely that \( \xi_j \) is correlated with the price so we need to non linear instrumental variable estimation techniques
Overview

- Continuous demand model
- Basic Discrete Choice model
- Beyond simple logit (Relaxing IIA)
  - Nested Logit
  - Mixed Logit/BLP
- Aggregation
- Endogeneity
- Applications
Endogenous price

- As with homogenous products an endogeneity concern about prices may arise
- Why? Because of unobserved product characteristics
  - What is the problem with unobserved characteristics?
- They are known to both, consumers and sellers, but not to the econometrician
- For example consider car demand:
  Product style may not be observed (also dealer quality, tradition may not be observed)
  How to explain Mini Cooper, BMW 300?
- Unobserved characteristics $\xi_k$ can be incorporated as suggested before by assuming that the utility of product $k$ is

\[
U_{ik} = x_{ik}' \beta_i - \alpha p_k + \xi_k + \varepsilon_{ik} = \delta_{ik} | \varepsilon_{ik}
\]
Endogenous price: Berry (1994)

- Non-linearity problem
  \( \xi_k \) enters market share \( s_k \) in a non-linear way and is correlated with prices
- Solution: True model implies a relationship between market share and mean utility \( \delta \)
  \[
  s_k = s_k (\delta; \alpha, \beta)
  \]

Can estimate demand parameters using IV from
  \[
  \delta_k = x_k' \beta - \alpha p_k + \xi_k
  \]

- Instruments?
  Excluded cost factors or characteristics of other firms
Endogenous Price

- **Cost factors**
  - Usually macro or at best city-specific (e.g. Wages). Need exogenous changes in brand-specific costs

- **Prices in Other cities (Hausman IV)**
  - Problem is that demand shocks also correlated across cities (e.g. Advertising)
  - Generally rejected when tested (Nevo, 2001; Genakos et al, 2010)

- **Characteristics of other firms**
  - Bresnahan et al, 1997 on PC market; Berry, 1994, BLP, 1995
  - Supply equation. Idea - when #products or quality increases in rival firms this reduces price
  - Assumes that product introduction and characteristics are exogenous. OK in short-run, but in long-run these will respond to changes in demand so are not exogenous.
Berry 1994

- There is a one to one correspondence between the mean utility associated with each alternative and its market share.

-
Implication

- It is simple to verify with the logit model that

\[
\frac{\log(s_j)}{\log(s_0)} = x_j \beta - \alpha p_j + \xi_j
\]

- The estimation requires a simple IV-LS to account for the price endogeneity
Nested logit

\[
\frac{\log(s_j)}{\log(s_0)} = x_j \beta - \alpha p_j + \xi_j + (1 - \lambda_k) \log(s_j|B_k)
\]

- The unobservable characteristic is possibly correlated with the price and with the log of the within group market share. We need at least two instruments to estimate the model
  - Price
  - Within group market share
Overview

- Continuous demand model
- Basic Discrete Choice model
- Beyond simple logit (Relaxing IIA)
  - Nested Logit
  - Mixed Logit/BLP
- Aggregation
- Endogeneity
- Applications
Some Applications - I

- Simulating merger effects (e.g. Nevo, 2000, RAND)
  - Have to assume a model of conduct (e.g. Bertrand Nash)
  - After estimating all demand parameters (“back end”) consider new allocation of assets of post merger
  - What will be the new vector of equilibrium prices (of merged firms and rivals)
  - By how much will prices increase?
  - This keeps marginal costs constant, but can also ask the question: by how much would marginal costs need to fall in order to offset increase in prices
Some Applications - II

- Testing models of market conduct (e.g. Nevo, 2001)
- Estimate the demand parameters
- Using these, estimate implied levels of profits/mark-ups under different market structures. Examples:
  1. Bertrand-Nash with existing brand ownership (i.e. Includes cannibalization by multi-product firms)
  2. Bertrand-Nash with all firms owning a single brand (i.e. No cannibalization)
  3. Fully collusive oligopoly
- Compare these hypothetical price cost margins with those estimated from other sources (e.g. Accounting data). Usually at industry-wide level
- Nevo found that only 1. was consistent with aggregate profit data
Summary

- Many ways to estimate brand-level elasticities
- Continuous demand
- Discrete Choice
  - Simple Logit
  - Nested Logit
  - Mixed Logit (Random coefficients)
- Useful across a range of areas – e.g. Mergers, conduct
- Common features that price endogeneity is key issue
- What can be done depends on data and computer power
Back Up slides
Technical note on aggregate price index, $P$

- In theory should be calculated as below, so we need

$$\log P = \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2} \sum_i \sum_k \gamma_{kj} \log p_k \log p_j$$
Technical note on aggregate price index, P

- So parameters estimated at different levels and then fed back into “ideal” price index.
- In practice, a “Stone” price index does a pretty good job where we use the brand shares as weights for all the individual prices.
Nested logit

- An alternative representation:

\[
\begin{align*}
\Pr_{ij} & = \Pr_{ij|B_k} \times \Pr_{B_k} \\
\Pr_{ij|B_k} & = \frac{\exp(\delta_{ij}/\lambda_k)}{\sum_{j \in B_k} \exp(\delta_{ij}/\lambda_k)} \\
\Pr_{B_k} & = \left(\frac{\sum_{j \in B_k} \exp(\delta_{ij}/\lambda_k)}{\sum_{l} \left(\sum_{j \in B_l} \exp(\delta_{ij}/\lambda_l)\right)^{\lambda_l}}\right)^{\lambda_k}
\end{align*}
\]
Estimation individual level data:

- MLE

\[ LL(\beta) = \sum_i \sum_j y_{ij} \log(Pr_{ij}) \]

- \( y_{ij} \) is an indicator function

- Goodness of fit

\[ \rho = 1 - \frac{LL(\hat{\beta})}{LL(0)}, \; \rho \epsilon [0, 1] \]
Notation

- $z_{ji}$ stands for both individual and product characteristics
- $x_j$ product characteristics

- $V_{ij} = z_{ij}\beta - ap_j + \xi_j$ is the mean utility involving both individual and product characteristics

- $\delta_j = x_j\beta - ap_j + \xi_j$ is the mean utility in product characteristics only

- Remember that $x$ and $z$ are bundles of characteristics so $\beta$ is a vector of parameters to estimate
Some Applications - III

- Assume Marginal cost

\[ mc_k = w_k \gamma + u_k \]

where \( w_k \) are cost shifters, and \( u_k \) is an unobserved cost error

- Profits

\[ \pi_f = \sum_{k \in J_f} [p_k - mc_k] \cdot M \cdot s_k (p, x, \tilde{\xi}; \alpha, \beta) \]

where \( M \) denotes the number of consumers in the market

- As before use necessary condition for eq prices

FOC

\[ s_k (p, x, \tilde{\xi}; \alpha, \beta) - \sum_{l \in J_f} [p_l - mc_l] \cdot \frac{\partial s_l (p, x, \tilde{\xi}; \alpha, \beta)}{\partial p_k} = 0 \]
Some Applications - III

- Given the demand function parameter estimates, many elements are determined:
  1. Can solve for the vector of mc’s
  2. Can solve for the markup $b_k (p, x, \xi; \alpha, \beta)$
  3. Can solve for $u_k$

- How to estimate from the FOC?

- Can think of estimating the equation

\[
mc_k = p_k - b_k (p, x, \xi; \alpha, \beta) = w_k \gamma + u_k
\]