Demand Estimation 1
PhD Industrial Organization

Nicholas Vreugdenhil

Why is estimating demand useful?

- Quantifying market power (think: inverse elasticity rule)
- Effects of a merger on prices
- Value of new goods
- Any question about consumer welfare
- Numerous other applications: e.g. school choice models, health insurance models, etc
- As a component of a larger model with a supply-side

Discrete choice demand models

- One of the key IO methods: **estimating demand in differentiated-product markets**.
- We will discuss the 'BLP' model: Berry, Levinsohn, and Pakes "Automobile Prices in Market Equilibrium" (1995)
  - Classic paper, and extremely influential.
  - Essentially, a 'mixed logit' model

Discrete choice demand models

- Many people consider the treatment in Nevo (2000) and Nevo (2001) to be more accessible if you are seeing the model for the first time.
  - I will closely follow the notation and content from HIO4 Chapter 2 (also written by Nevo along with Gandhi)
  - HIO4 Chapter 1 also covers demand estimation, but more from the perspective of thinking deeply about identification. In the interests of time, I will focus a little less on this chapter.
- As well as being an extremely useful method in its own right, understanding the BLP model will help you to learn skills more broadly useful in structural modeling work (e.g. computational simulation, dealing with endogeneity, etc).

Plan for today

1. General setup of the BLP model

2. Price elasticity/substitution patterns

Plan for today

1. **General setup of the BLP model**
2. Price elasticity/substitution patterns

Discrete choice demand models: general setup of ‘BLP’ (‘mixed logit’)

Berry, Levinsohn, and Pakes (1995)

$$u_{ijt} = x_{jt}\beta_{it} + \alpha_{it}p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- i: denotes consumer; t: denotes market; j: denotes product
- u_{ijt}: utility of consumer i for product j in market t
- x_{jt}: vector of observed product characteristics
- p_{jt}: price of product j in market t
- ξ_{jt}: unobserved demand shock/product characteristics (observed by consumers and firms but not the econometrician)
- ϵ_{ijt}: idiosyncratic taste shock (typically i.i.d. across (i,j,t) and drawn from type-1 extreme value distribution with scale parameter normalized to 1)

Discrete choice demand models: general setup of 'BLP'

Berry, Levinsohn, and Pakes (1995)

$$u_{ijt} = x_{jt}\beta_{it} + \alpha_{it} p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- Why is it useful to write the error term in this particular way ($\xi_{jt} + \epsilon_{ijt}$)?
- Later, we will see writing the error term like this is helpful when estimating the model using market-level data.
  - Idea: ϵ_{ijt} do not impact pricing, but firms observe ξ_{jt} for all products j when setting prices.
  - Therefore, ξ_{jt} will typically be the econometric error term ($+$ prices will be correlated with it).
  - Notice that we are being very explicit about the problem of price endogeneity in the model.
  - This will allow us later to be explicit about how we solve it (spoiler: we will use instrumental variables).

Discrete choice demand models: general setup of ‘BLP’

Berry, Levinsohn, and Pakes (1995)

$$u_{ijt} = x_{jt}\beta_{it} + \alpha_{it} p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- More on ξ_{jt}:
- These unobservables do not vary within a market. But what are the boundaries of a ‘market’?
    - Geography: city? state? country?
    - Time: week? month? day?
    - Answers depend on the specific industry details and data availability - needs to be thought about carefully.

Discrete choice demand models: general setup of 'BLP'

Berry, Levinsohn, and Pakes (1995)

$$u_{ijt} = x_{jt}\beta_{it} + \alpha_{it}p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- More on α_{it} (price sensitivity):

$$\alpha_{it} = \alpha_0 + \sum_{l=1}^{L} \alpha_l D_{ilt} + \alpha_\nu \nu_{it}^{(0)}$$

- D_{ilt}: L 'demographic' variables (e.g. income, age, family size). This is either observed or the distribution is known.
  - E.g. get distribution of income, education, etc from Current Population Survey in different cities in the US.

- $\nu_{it}^{(0)}$: random variable
  - Captures variation in preferences beyond what the demographic variables can explain (for cars, e.g. dog ownership is typically not observed but potentially a factor in car purchase)

Discrete choice demand models: general setup of 'BLP'

Berry, Levinsohn, and Pakes (1995)

$$u_{ijt} = x_{jt}\beta_{it} + \alpha_{it}p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- More on β_{it} (product characteristics sensitivity):
- Denote $\beta_{it}^{(k)}$ the weight consumer i places on characteristic $x_{jt}^{(k)}$.

$$\beta_{it}^{(k)} = \beta_0^{(k)} + \sum_{l=1}^{L} \beta_d^{(l,k)} D_{ilt} + \beta_\nu^{(k)} \nu_{it}^{(k)}$$

- $\beta_0^{(k)}$: parameter common to all consumers
- D_{ilt}: L 'demographic' variables (e.g. income, age, family size)
- $\nu_{it}^{(k)}$: random variable

Discrete choice demand models: general setup of ‘BLP’

Berry, Levinsohn, and Pakes (1995)

$$u_{ijt} = x_{jt}\beta_{it} + \alpha_{it} p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- Outside good j=0:
- Consumers have the option to not purchase any of the J products.
- Indirect utility:

$$u_{i0t} = \epsilon_{i0t}$$

- In above, non-idiosyncratic components are normalized to 0.
- So, utility of the other J goods should be interpreted as relative to this outside good.