Demand Estimation 2
PhD Industrial Organization

Nicholas Vreugdenhil

Plan

1. BLP setup
2. Price elasticity/substitution patterns
3. Estimation: overview and typical data
4. Identification: what if we had micro-data?

Plan

1. **BLP setup**
2. Price elasticity/substitution patterns
3. Estimation: overview and typical data
4. Identification: what if we had micro-data?

Discrete choice demand models: general setup of 'BLP'

Berry, Levinsohn, and Pakes (1995)

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

- What about income?:
- In the above equation, p_{jt} should really be $y_i - p_{jt}$ where y_i is income.
- Leaving it out has no impact for choices however, and just simplifies exposition. Why?

Discrete choice demand models: general setup of 'BLP'

Berry, Levinsohn, and Pakes (1995)

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

- <u>What about income?</u>:
- In the above equation, p_{jt} should really be $y_i - p_{jt}$ where y_i is income.
- Leaving it out has no impact for choices however, and just simplifies exposition. Why?
  - Income enters linearly to all options, and only relative differences in utilities matter for choice probabilities
  - If $y_i - p_{jt}$ entered non-linearly then it would affect the choice probabilities and should be included

Discrete choice demand models: general setup of 'BLP'

Berry, Levinsohn, and Pakes (1995)

- Even more notation:
  - Recall L number of demographic vars, K number of product characteristics
- Define:
  - The **mean utility** of product j in market t: $\delta_{jt} = x_{jt}\beta_0 + \alpha_0 p_{jt} + \xi_{jt}$
  - Γ: $(K+1) \times L$ matrix with coefficients of demographic variables
  - Σ: $(K+1) \times (K+1)$ diagonal matrix with diagonal $(\alpha_\nu, \beta_\nu^{(1)}, ..., \beta_\nu^{(K)})$
  - $\nu_{it} = (\nu_{it}^{(0)}, ..., \nu_{it}^{(K)})^T$
  - $\mu_{ijt} = (x_{jt}, p_{jt}) \cdot (\Gamma D_{it} + \Sigma \nu_{it})$
- Then we can rewrite our utility equation as:

$$u_{ijt} = \underbrace{\delta_{jt}}_{\text{mean utility}} + \underbrace{\mu_{ijt}}_{\text{interaction between consumer tastes + product characteristics}} + \underbrace{\epsilon_{ijt}}_{\text{idiosyncratic error}}$$

Discrete choice demand models: general setup of 'BLP'

Berry, Levinsohn, and Pakes (1995)

- Review of where we are: we just characterized a very flexible model of consumer utility.
- Assuming i.i.d. extreme value errors ϵ_{ijt} the probability consumer i chooses product j is:

$$\frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{J} \exp(\delta_{kt} + \mu_{ikt})}$$

- And **demand** (the share of consumers who purchase good j in market t) is:

$$s_{jt} = \sigma_j(\delta_t, \boldsymbol{x}_t, \boldsymbol{p}_t; \Gamma, \Sigma) = \int \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{J} \exp(\delta_{kt} + \mu_{ikt})} dF(D_{it}, v_{it})$$

- Here:
  - $\delta_t, \boldsymbol{x}_t, \boldsymbol{p}_t$ are vectors of mean utilities, observed product characteristics, prices, in market t
  - F is the joint distribution of observed demographics and unobserved tastes

Plan

1. BLP setup
2. **Price elasticity/substitution patterns**
3. Estimation: overview and typical data
4. Identification: what if we had micro-data?

Price elasticity/substitution patterns

- **Question**: is all the complexity in the previous section necessary (in terms of heterogeneous consumers etc)?
  - What would a simpler model (for example, with homogeneous consumers) fail to capture?
- **Answer**: (Typically) it is!
  - Key implication of a demand model: substitution patterns between goods/price elasticity
  - I will now argue that allowing for flexible consumer heterogeneity is **necessary to get the model to generate realistic substitution patterns**.

Price elasticity/substitution patterns: implications of homogeneous consumer model

- **Thought experiment**: switch off consumer heterogeneity.
  - E.g. accomplish this by setting $\Gamma = 0$ and $\Sigma = 0$. So, $\mu_{ijt} = 0$.

- Then, just a Logit model:

$$s_{jt} = \frac{\exp(\delta_{jt})}{1 + \sum_{k=1}^{J} \exp(\delta_{kt})}$$

- Price elasticities:

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \alpha_0 p_{jt}(1 - s_{jt}) & \text{if j=k} \\ -\alpha_0 p_{kt} s_{kt} & \text{otherwise} \end{cases}$$

Price elasticity/substitution patterns: implications of homogeneous consumer model

- Price elasticities:

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \alpha_0 p_{jt}(1 - s_{jt}) & \text{if j=k} \\ -\alpha_0 p_{kt} s_{kt} & \text{otherwise} \end{cases}$$

- **Implication 1**:
  - Typically, $\alpha_0(1 - s_{jt}) \approx \alpha_0$ since there are many products and market share of each product is small.
  - So, own price-elasticities (j=k) are proportional to price.
  - What does the model imply about prices vs demand elasticity?
    - This demand model implies that the **lower the price, the more inelastic is demand**
    - Further implication: under typical pricing models $\rightarrow$ higher markup for these lower priced goods
    - Question: do you think that these implications are reasonable predictions for the model to make?

Price elasticity/substitution patterns: implications of homogeneous consumer model

- Price elasticities:

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \alpha_0 p_{jt}(1 - s_{jt}) & \text{if j=k} \\ -\alpha_0 p_{kt} s_{kt} & \text{otherwise} \end{cases}$$

- **Implication 2**:
  - Consider an increase in the price of product k. Concretely, think about the market for cars. The price of a BMW goes up. Do you think consumers will substitute towards a Mercedes or a Honda Civic?

Price elasticity/substitution patterns: implications of homogeneous consumer model

- Price elasticities:

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \alpha_0 p_{jt}(1 - s_{jt}) & \text{if j=k} \\ -\alpha_0 p_{kt} s_{kt} & \text{otherwise} \end{cases}$$

- **Implication 2**:
  - Consider an increase in the price of product k. Concretely, think about the market for cars. The price of a BMW goes up. Do you think consumers will substitute towards a Mercedes or a Honda Civic?
    - Usually, we'd expect consumers to substitute towards similar products (i.e. the Mercedes)
  - But, the homogeneous consumer model predicts the following **diversion ratio**:

$$\frac{\partial s_{jt}}{\partial p_{kt}} / \frac{\partial s_{kt}}{\partial p_{kt}} = s_{jt}/(1 - s_{kt})$$

  - Here, substitution is proportional to market share, not how close the products are in terms of their characteristics.
    - Idea: as p_k increases, consumers who no longer choose k choose other options at the same frequency as the 'average' consumer (i.e. in proportion to their market share).

Price elasticity/substitution patterns

- Price elasticities in the full BLP model (which heterogeneous consumers):

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \frac{p_{jt}}{s_{jt}} \int \alpha_{it} s_{ijt}(1 - s_{ijt}) dF(D_{it}, \nu_{it}) & \text{if j=k} \\ -\frac{p_{kt}}{s_{jt}} \int \alpha_{it} s_{ijt} s_{ikt} dF(D_{it}, \nu_{it}) & \text{otherwise} \end{cases}$$

- Notation: s_{ijt}: probability that i purchases j in market t

- **Observation 1**:

- Each consumer has a different price sensitivity, which is averaged to a product-specific mean price sensitivity using the individual probabilities of purchase as weights.

- This relaxes 'implication 1' from before. I.e. model could generate that low-price products have more elastic demand

Price elasticity/substitution patterns

- Price elasticities in the full BLP model (i.e. including heterogeneous consumers):

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \frac{p_{jt}}{s_{jt}} \int \alpha_{it} s_{ijt}(1 - s_{ijt}) dF(D_{it}, \nu_{it}) & \text{if j=k} \\ -\frac{p_{kt}}{s_{jt}} \int \alpha_{it} s_{ijt} s_{ikt} dF(D_{it}, \nu_{it}) & \text{otherwise} \end{cases}$$

- Notation: s_{ijt}: probability that i purchases j in market t

- **Observation 2**:

- Model generates flexible cross-product substitution patterns.
  - How? Correlation in μ_{ijt} and μ_{ikt} induces correlation between s_{ijt} and s_{ikt}, which then determines substitution patterns.

- Note: alternatively, may be able to generate realistic substitution patterns with a nested logit (e.g. put the luxury cars in the same nest)
  - ...but this requires a-priori decisions about how to segment the market.

Plan

1. BLP setup
2. Price elasticity/substitution patterns
3. **Estimation: overview and typical data**
4. Identification: what if we had micro-data?

Estimation: setup of the problem

- What are the parameters we need to estimate?

Estimation: setup of the problem

- What are the parameters we need to estimate?
- Linear parameters:
    - Parameters from the mean utility equation: (α_0, β_0)
- Nonlinear parameters
    - Γ: coefficients on (observed) demographics
    - Σ: idiosyncratic "taste for characteristics"
- So, full parameter vector to estimate: $\theta = (\alpha_0, \beta_0, \Gamma, \Sigma)$.

Estimation: setup of the problem

- Common assumptions in empirical work:
  - We will also make these assumptions from now on.
  - They simplify the model, but are not necessary, see Chapter 1 of the Handbook for ways to relax these assumptions.
  - (Always good to know common assumptions people make in empirical work, especially when they simplify the model!)
- 1. Distribution of “taste for characteristics” $\nu_{it} = (\nu_{it}^{(0)}, ..., \nu_{it}^{(K)})$ is independent of the distribution of demographics D_{it}.
  - Then, $F(D_{it}, \nu_{it}) = F_D(D_{it}) F_\nu(\nu_{it})$.
- 2. Each $v_{it}^{(k)}$ is independent across $k = 0, ..., K$ and distributed standard normal.

Data: typically, the data have three types of variables

- 1. Quantities of the J products purchased in market t.
  - These are aggregations of individual consumer choices.
  - As we will see, only aggregate data (ie. data on total quantities at the market level) is required for identification. However, information from micro-data (i.e. data on individual choices) can be incorporated.
  - Remember: we are implicit assuming that definition of a market is narrow enough that consumers in the market face the same prices, characteristics, and demand shocks.
  - Can convert aggregate quantities to market shares if we know the total market size I_t: $s_{jt} = q_{jt} / I_t$.

Data: typically, the data have three types of variables

-
  2. Prices p_{jt} and "observed" product characteristics x_{jt} of each of J products in market t.

- 3. Information on consumer demographics.
  - In micro-data, will observe D_{ilt} i.e. demographics for each i
    - e.g. survey data on car purchases
  - In other applications, more aggregated data
    - e.g. *distribution* of demographics $F_t(D)$
    - could obtain such data from e.g. the Current Population Survey in different cities in the US)
  - In other applications, data at a granularity somewhere between the two above cases.
    - e.g. average age of consumers who purchase product j

Plan

1. BLP setup
2. Price elasticity/substitution patterns
3. Estimation: overview and typical data
4. **Identification: what if we had micro-data?**

Identification

- What variation in the data can identify the parameters?
  - Precise econometric definition of identification: see haile.pdf on Canvas, or 'The Identification Zoo' (Lewbel)
- **Thought experiment**: what if we:
  - 1. have micro-data on individual consumers
  - 2. observe a single market
  - 3. switch off $\Sigma = 0$ (i.e. ignore any idiosyncratic "taste for characteristics", implies heterogeneity is only driven by observed demographics)
- Later, we will build on this intuition to discuss what to do if we had more aggregated market-level data with random taste shocks etc...

Identification using individual-level data

- Data:
- $\{y_{ij}, D_i\}_{i=1,...,I}$ where $y_{ij} = 1$ for $j = 0, 1, ..., J$ if consumer i chooses product j and $\sum_j y_{ij} = 1$.
- All consumers are from the same market (same prices, same product characteristics both observed $\mathbf{x}$ and unobserved ξ)
- Comment:
- Estimating demand might seem hopeless here: we only see one market, so how are we supposed to get how quantities vary with prices if there is no price variation in the data?
- But, we will now see that it is in fact possible.

Identification using individual-level data

- (Conditional indirect) utility from product j (dropping t subscript and incorporating price p_j as a 'characteristic' in x_j to simplify exposition):

$$u_{ij} = \underbrace{x_j\beta_0 + \xi_j}_{\delta_j} + \sum_{k,l} \beta_d^{(l,k)} D_{il} x_{jk} + \epsilon_{ij}$$

- Comment: if we didn't have the (unobserved) demand shock ξ_j then we could estimate all the parameters of the model at once by maximum likelihood.

Identification using individual-level data

- (Conditional indirect) utility from product j (dropping t subscript and incorporating price p_j as a 'characteristic' in x_j to simplify exposition):

$$u_{ij} = \underbrace{x_j\beta_0 + \xi_j}_{\delta_j} + \sum_{k,l} \beta_d^{(l,k)} D_{il} x_{jk} + \epsilon_{ij}$$

- Instead, use a **two-step procedure**:

- 1. Include a product-specific intercept to capture $\delta_j = x_j\beta_0 + \xi_j$ (i.e. estimate $\tilde{\theta} = (\delta_1, ..., \delta_J, \Gamma)$ using maximum likelihood)

- 2. Estimate β_0 by 'projecting' estimated δ's on the x's.
    - If assume $E(\xi_j | x_j) = 0$ then can use (weighted) least squares
    - If concerned x's are correlated with ξ can use $E(\xi_j | Z_j) = 0$ where Z are a vector of exogeneous variables (discuss more in a few slides...)

Identification using individual-level data: step 1

- Estimate the δ and Γ parameters by maximum likelihood with utility:

$$u_{ij} = \delta_j + \sum_{k,l} \beta_d^{(l,k)} D_{il} x_{jk} + \epsilon_{ij}$$

- Identifying δ_j:

- Take FOC of likelihood $\rightarrow$ can show that intercepts δ_j are found by setting observed market shares equal to the ones predicted by model. That is, if for a fixed Γ, set:

$$\hat{s}_j = \hat{\sigma}(\hat{\delta}_1, ..., \hat{\delta}_J)$$

- Under some general technical conditions can invert this relationship (the 'Berry inversion'):

$$\hat{\delta}_j = \hat{\sigma}_j^{-1}(\hat{s}_1, ..., \hat{s}_J)$$

- Asymptotically as $I \rightarrow \infty$:

$$\delta_j = \sigma_j^{-1}(s_1, ..., s_J)$$

Identification using individual-level data: step 1

- To be continued...