UG Empirical Industrial Organization

Lecture 7: Demand Systems: Discrete Choice Models

Anubhav Jha[1]
Ashoka University

March 11, 2024

[1]This course is based on Victor Aguirregabiria's Empirical IO book. The slides (in my first year of teaching) are extremely close to his slides. These will change in the future iterations of this course.

Outline on today's lecture

1. Estimation of the Standard Logit Model
   1.1. Endogeneity problem & bias of OLS estimator
   1.2. Instrumental Variables estimation
2. Logit model with heterogeneous coefficients

Estimation: Standard Logit Model Data

- Suppose that we have data on quantities (sold), prices, and characteristics of all the J products in a market:

$$\text{Data} = \{q_j, p_j, X_{1j}, \ldots, X_{Kj} : j = 1, 2, \ldots, J\}$$

- Suppose that we also observe the consumers who have not purchased any of the J products, i.e., q_0.
- For instance, the Stata dataset `verboven_cars.dta` contains the following variables for $J = 356$ car models in the markets of five different European countries:
  - `price`, `quantity`, `brand`
  - `displacement` (in cc), `horsepower` (in kW), `weight` (in kg)
  - `seats`, `doors`, `length` (in cm), `width` (in cm), `height` (in cm)
  - `fuel efficiency` (liters per km), `maximum speed` (km/h), `acceleration time` (seconds from 0 to 100 km/h)

Estimation: Logit Model

- Given quantities, we can construct market shares.
- Market size (number of consumers) is:

$$H = q_0 + q_1 + \cdots + q_J$$

- The market share of product j is:

$$s_j = \frac{q_j}{H}$$

- The logit model implies the following regression model:

$$y_j = \beta^p p_j + \beta_1 X_{1j} + \cdots + \beta_K X_{Kj} + \xi_j$$

  where:

$$y_j = \ln(s_j) - \ln(s_0), \quad \beta^p = -\alpha$$

- The error term ξ_j represents characteristics of product j that are valuable to consumers but unobservable to us as researchers.
- Given these data, we can estimate the parameters $(\beta^p, \beta_1, \ldots, \beta_K)$.

OLS Estimation: Endogeneity problem

- Unfortunately, the OLS estimator does not provide unbiased (consistent) estimates of the parameters of the model.
- Products with higher unobserved quality ξ_j tend to have higher prices:

$$\text{cov}(p_j, \xi_j) > 0$$

- This endogeneity problem implies that the OLS estimate $\hat{\beta}_p^{\text{OLS}}$ captures two effects:
  - a the **causal effect** of price on y_j, i.e., $\beta_p < 0$;
  - b an **indirect positive effect** (not causal) from the correlation between price and unobserved product quality.
- Algebraically:

$$\hat{\beta}_p^{\text{OLS}} = \frac{\text{cov}(y_j, p_j)}{\text{var}(p_j)} = \beta_p + \frac{\text{cov}(p_j, \xi_j)}{\text{var}(p_j)}$$

- We could even observe $\hat{\beta}_p^{\text{OLS}} > 0$, despite the true causal effect $\beta_p < 0$.

Endogeneity problem: Example

- Suppose that the profit maximization condition (Marginal Revenue = Marginal Cost) implies the following optimal price for the firm selling product j:

$$p_j = \gamma_1 X_{1j} + \cdots + \gamma_K X_{Kj} + \gamma_\xi \xi_j$$

  where the γ's are parameters.
- Product characteristics affect price because:
  1. They affect marginal costs (e.g., higher quality products are costlier to produce);
  2. They enter demand and affect marginal revenue.
- The model consists of:
  - The logit demand equation:

$$y_j = \beta^p p_j + \xi_j$$

  - The pricing equation:

$$p_j = \gamma_\xi \xi_j$$

    (omitting X-variables for simplicity).

Endogeneity problem: Example [2]

- Consider the structural equations of the model (omitting constant terms; all variables are in deviations from their means):

$$y_j = \beta^p p_j + \xi_j, \quad p_j = \gamma_\xi \xi_j$$

- Substituting the pricing equation into the demand equation:

$$y_j = \beta^p(\gamma_\xi \xi_j) + \xi_j = (\beta^p \gamma_\xi + 1)\xi_j$$

- Therefore, the covariance between y_j and p_j is:

$$\text{cov}(y_j, p_j) = (\beta^p \gamma_\xi + 1)\gamma_\xi \, \text{var}(\xi_j)$$

- And the variance of p_j is:

$$\text{var}(p_j) = \gamma_\xi^2 \, \text{var}(\xi_j)$$

Endogeneity problem: Example [3]

- In this model, the OLS estimator is:

$$\hat{\beta}_p^{\text{OLS}} \rightarrow \frac{\text{cov}(y_j, p_j)}{\text{var}(p_j)} = \frac{(\beta^p \gamma_\xi + 1)\gamma_\xi \, \text{var}(\xi_j)}{\gamma_\xi^2 \, \text{var}(\xi_j)} = \beta^p + \frac{1}{\gamma_\xi}$$

- Since $\frac{1}{\gamma_\xi} > 0$, the OLS estimator is an upward-biased estimate of the true β^p.
- Given that $\beta^p < 0$, the estimate is biased toward zero—i.e., it is less negative than the truth—or may even be positive.

Instrumental Variables (IV) Estimation

- To address the endogeneity problem, we can use IV estimation.
- We need a variable (or a set of variables) Z_j that satisfies the following three conditions:
  - [1] **Exclusion.** Z_j is **not** an explanatory variable in the demand equation of product j; i.e., $Z_j \notin X_j$.
  - [2] **No correlation with the error term.** Z_j is **not correlated** with the unobserved quality ξ_j of product j.
  - [3] **Relevance.** In a regression of price p_j on X_j and Z_j, the variable Z_j has a statistically significant (partial) correlation with p_j.
- Even if we cure the endogeneity problem, the standard multinomial logit model is very restrictive in terms of elasticities it can support.
- Recall : $\frac{ds_j}{dp_j} \cdot \frac{p_j}{s_j} = -\alpha(1 - s_j)p_j$ and $\frac{ds_j}{dp_i} \cdot \frac{p_i}{s_j} = -\alpha s_i p_i$

Why Are These Elasticities Very Restrictive?

- The standard Logit model imposes strong restrictions on substitution patterns implied by cross-price elasticities.
- Consider three car models:

$$\texttt{ECON1}, \quad \texttt{ECON2}, \quad \texttt{LUX}$$

  with market shares:

$$s_{\text{ECON1}} = 0.20, \quad s_{\text{ECON2}} = 0.20, \quad s_{\text{LUX}} = 0.01$$

  and prices:

$$p_{\text{ECON1}} = 1, \quad p_{\text{ECON2}} = 1, \quad p_{\text{LUX}} = 20$$

- Now consider the effect of increasing p_{ECON1} on the demands for ECON2 and LUX.
- According to the logit model:

$$\frac{ds_{\text{ECON2}}}{dp_{\text{ECON1}}} \cdot \frac{p_{\text{ECON1}}}{s_{\text{ECON2}}} = \alpha \cdot s_{\text{ECON1}} \cdot p_{\text{ECON1}} = \alpha \cdot 0.20$$

$$\frac{ds_{\text{LUX}}}{dp_{\text{ECON1}}} \cdot \frac{p_{\text{ECON1}}}{s_{\text{LUX}}} = \alpha \cdot s_{\text{ECON1}} \cdot p_{\text{ECON1}} = \alpha \cdot 0.20$$

- **Implication:** A price increase for ECON1 leads to the same proportional increase in demand for ECON2 as for a luxury car—clearly unrealistic.

Dealing with Limitations of Standard Logit Model

- The standard Logit model imposes restrictive substitution patterns due to its i.i.d. Extreme Value errors and homogeneous coefficients.
- We can relax these restrictions through two key model extensions:
  1. **Consumer heterogeneous coefficients** β_h:
     Allow each consumer to have their own preferences for product characteristics. This is implemented in the Random Coefficients Logit model.
  2. **Nested Logit model for** ε_h:
     Allow correlation in unobserved utility shocks across similar products by grouping them into "nests."

Discrete Choice with Heterogeneous Coefficients

- The utility of consumer i if they purchase product j and reside in market m is:

$$U_{ijm} = -\alpha_{im} \cdot p_{jm} + X_{jm} \cdot \beta_{im} + \xi_{jm} + \varepsilon_{ijm}$$

- $X_{jm}\,\beta_{im} = \beta_{1im}X_{1jm} + \beta_{2im}X_{2jm} + \cdots + \beta_{Kim}X_{Kjm}$ is the utility from the observable characteristics.
- β_{kim}: represents the marginal utility of this product characteristic (e.g., RAM memory) for consumer i residing in m.
- ξ_{jm}: Utility from unobservable characteristics, the same for all consumers in market m.
- ε_{ijm}: The T1EV idiosyncratic taste shock of consumer i for product j in m.

Discrete Choice with Heterogeneous Coefficients

- The utility of consumer i if they purchase product j and reside in market m is:

$$U_{ijm} = -\alpha_{im} \cdot p_{jm} + X_{jm} \cdot \beta_{im} + \xi_{jm} + \varepsilon_{ijm}$$

- Generally the following forms are assumed for $(\alpha_{im}, \beta_{im})$:
  1. BLP-1995

$$\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \nu_i \qquad \text{s.t.} \qquad \nu_i \; N(0, \Sigma) \tag{1}$$

  2. BLP-1995 (with market variation)

$$\begin{bmatrix} \alpha_{im} \\ \beta_{im} \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \nu_{im} \qquad \text{s.t.} \qquad \nu_{im} \; N(0, \Sigma) \tag{2}$$

- We cannot use m fixed effects because we only observe market-level shares. So means are still for all markets.

Discrete Choice with Heterogeneous Coefficients

- The utility of consumer i if they purchase product j and reside in market m is:

$$U_{ijm} = -\alpha_{im} \cdot p_{jm} + X_{jm} \cdot \beta_{im} + \xi_{jm} + \varepsilon_{ijm}$$

- Generally the following forms are assumed for $(\alpha_{im}, \beta_{im})$:

    1. General case (with market variation)

$$\begin{bmatrix} \alpha_{im} \\ \beta_{im} \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \sum_{r=1}^{R} \Pi^{r}_{(K+1)} D_{rim} + \nu_{im} \qquad \text{s.t.} \qquad \nu_{im} \;\; N(0, \Sigma) \tag{3}$$

    where $D_{im} = (D_{1im}, \ldots, D_{Rim})$ is $R \times 1$ dimensional vector of socio-economic characteristics of consumer i in market m. $\Pi^{r}_{(K+1)}$ is a column vector of parameters. Π^{r}_{1+k} a consumer's marginal utility from characteristic k changes as their socio-economic characteristic D_{rim} changes.

Discrete Choice with Heterogeneous Coefficients

- The utility of consumer i if they purchase product j and reside in market m is:

$$U_{ijm} = -\alpha_{im} \cdot p_{jm} + X_{jm} \cdot \beta_{im} + \xi_{jm} + \varepsilon_{ijm}$$

General case gives us market level means in taste parameters (with market variation).

$$\begin{bmatrix} \alpha_{im} \\ \beta_{im} \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \sum_{r=1}^{R} \Pi^r_{(K+1)} D_{rim} + \nu_{im} \tag{4}$$

$$\Rightarrow \begin{bmatrix} \alpha_{im} \\ \beta_{im} \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \sum_{r=1}^{R} \Pi^r_{(K+1)} \bar{D}_{rm}}_{\text{Market level mean}} + \underbrace{\sum_{r=1}^{R} \Pi^r_{(K+1)} \left(D_{rim} - \bar{D}_{rm} \right) + \nu_{im}}_{\text{Within-market individual heterogeneity}} \tag{5}$$

$$\Rightarrow \begin{bmatrix} \alpha_{im} \\ \beta_{im} \end{bmatrix} = \begin{bmatrix} \alpha_m \\ \beta_m \end{bmatrix} + \tilde{\nu}_{im} \tag{6}$$

Discrete Choice with Heterogeneous Coefficients

- The utility of consumer i if they purchase product j and reside in market m is:

$$U_{ijm} = -\alpha_{im} \cdot p_{jm} + X_{jm} \cdot \beta_{im} + \xi_{jm} + \varepsilon_{ijm}$$

- The utility of consumer i can be re-written as followed:

$$U_{ijm} = \underbrace{-\alpha_m \cdot p_{jm} + X_{jm} \cdot \beta_m + \xi_{jm}}_{\delta_{jm}\text{: Product } j \text{ mean utility in } m} + \underbrace{\begin{bmatrix} -p_{jm} \\ X_{jm}^T \end{bmatrix}^T \cdot \tilde{\nu}_{im} + \varepsilon_{ijm}}_{\mu_{ijm}+\varepsilon_{ijm}\text{: Product } j \text{ idiosyncratic utility}}$$

$$U_{ijm} = \delta_{jm} + \mu_{ijm} + \varepsilon_{ijm}$$

Discrete Choice with Heterogeneous Coefficients

- Market Shares under Heterogeneous Coefficients:

$$s_{jm} = \mathbb{E}_{\mu_{im},\epsilon_{ijm}} \left[1 \left\{ U_{ijm} \geq U_{ij'm} \right\} \right] \tag{7}$$

$$\Rightarrow s_{jm} = \mathbb{E}_{\mu_{im}} \left[\mathbb{E}_{\epsilon_{ijm}} \left[1 \left\{ U_{ijm} \geq U_{ij'm} \right\} \middle| \mu_{im} \right] \right] \tag{8}$$

$$\Rightarrow s_{jm} = \mathbb{E}_{\mu_{im}} \left[\overbrace{\frac{\exp\left(\delta_{jm} + \mu_{ijm}\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_{km} + \mu_{ikm}\right)}}^{s_{ijm}: i\text{' probability of purchasing } j} \right] \tag{9}$$

$$\Rightarrow s_{jm} = \int_{\mu_{im}} \frac{\exp\left(\delta_{jm} + \mu_{ijm}\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_{km} + \mu_{ikm}\right)} dF\left(\mu_{im}\right) \tag{10}$$

where $\mu_{im} = \left(\mu_{i1m}, \mu_{i2m}, \ldots, \mu_{iJm}\right)$

Discrete Choice with Heterogeneous Coefficients

- Cross-price elasticity under hetereogeneous Coefficients:

$$\frac{ds_{jm}}{dp_{j'm}} \cdot \frac{p_{j'm}}{s_{jm}} = \frac{d\mathbb{E}_{\mu_{im}}\left[\frac{\exp\left(\delta_{jm}+\mu_{ijm}\right)}{1+\sum_{k=1}^{K}\exp(\delta_{km}+\mu_{ikm})}\right]}{dp_{jm}} \cdot \frac{p_{j'm}}{s_{jm}} \tag{11}$$

$$\Rightarrow \frac{ds_{jm}}{dp_{j'm}} \cdot \frac{p_{j'm}}{s_{jm}} = \mathbb{E}_{\tilde{\nu}_{im}}\left[\frac{d\frac{\exp\left(\delta_{jm}-p_{jm}\cdot\tilde{\nu}_{1,im}+\sum_{l=1}^{K}X_{ljm}\cdot\tilde{\nu}_{l+1,im}\right)}{1+\sum_{k=1}^{J}\exp\left(\delta_{km}-p_{km}\cdot\tilde{\nu}_{l,im}+\sum_{l=1}^{K}X_{lkm}\cdot\tilde{\nu}_{l+1,im}\right)}}{dp_{j'm}}\right] \cdot \frac{p_{j'm}}{s_{jm}} \tag{12}$$

$$\Rightarrow \frac{ds_{jm}}{dp_{j'm}} \cdot \frac{s_{jm}}{p_{j'm}} = \mathbb{E}_{\tilde{\nu}_{im}}\left[s_{ijm} \cdot s_{ij'm} \cdot (\alpha_m + \tilde{\nu}_{1im})\right] \cdot \frac{p_{j'm}}{s_{jm}} \tag{13}$$

$$= \left(\alpha_m \mathbb{E}_{\tilde{\nu}_{im}}\left[s_{ijm} \cdot s_{ij'm}\right] + \mathbb{E}_{\tilde{\nu}_{im}}\left[s_{ijm} \cdot s_{ij'm} \cdot \tilde{\nu}_{1im}\right]\right) \cdot \frac{p_{j'm}}{s_{jm}} \tag{14}$$

$$\tag{15}$$

- The above expression provides more flexible cross-price elasticities by taking account of how close substitutes different products are. The expectations above are essentially flexible measure of the degree of substitution between two products j and j' across the population of consumers in market m.

Discrete Choice with Heterogeneous Coefficients: Estimation

- The data typically used for discrete choice with Heterogeneous Coefficients or BLP is given by:
  - BLP 1995:

$$\text{Data} = \{S_j, p_j, X_{1j}, \ldots, X_{Kj}, Z_{1j}, \ldots, Z_{B,j} : j = 1, 2, \ldots, J\}$$

    where $B \geq$ number of parameters to be estimated.
  - BLP with market level variation:

$$\text{Data} = \Big\{S_{jm}, p_{jm}, X_{1jm}, \ldots, X_{Kjm}, Z_{1jm}, \ldots, Z_{B,jm} : \\ j = 1, 2, \ldots, J, m = 1, 2, \ldots, M\Big\}$$

  - General BLP:

$$\begin{aligned}\text{Data} = \Big\{&S_{jm}, p_{jm}, X_{1jm}, \ldots, X_{Kjm}, \\ &\bar{D}_{1m}, \bar{D}_{2m}, \ldots, \bar{D}_{Rm}, \Omega_{D_m D_m, m}, \\ &Z_{1jm}, \ldots, Z_{B,jm} : \\ &j = 1, 2, \ldots, J, m = 1, 2, \ldots, M\Big\}\end{aligned}$$

    where $\bar{D}_{1m}, \bar{D}_{2m}, \ldots, \bar{D}_{Rm}, \Omega_{D_m D_m, m}$ allows us to compute $\mathbb{E}_{\mu_{ijm}}$ for each market. These capture the distribution of socio-economic characteristics across consumers within a market.

Discrete Choice with Heterogeneous Coefficients: Estimation

- Estimation: To estimate the model we target the moment condition $E\left[\xi_{jm}|Z_{jm}\right]=0$, where Z_{jm} is a vector of instrument variables such that the number of instruments $\geq$ is higher than the number of parameters i.e. the number of elements in $\alpha_0, \beta_0, \Pi, \Sigma$.
- Note that $\xi_{jm}=\delta_{jm}+\alpha_m \cdot p_{jm}-X_{jm}\cdot\beta_m$.
- Same steps $logs_j/logs_0$ as before do not allow us to recover δ_{jm} due to the non-linear structure of the market share function.
- Berry, Levinsohn, and Pakes (1995) showed that the following iterative steps converge to δ_{jm} for $l=1,2,\ldots$

$$\delta_{jm}^{l+1}=\delta_{jm}^{l}+\log(S_{jm})-\log(s_{jm}) \tag{16}$$

$$\Rightarrow \delta_{jm}^{l+1}=\delta_{jm}^{l}+\log(S_{jm})-\log\left(\int_{\mu_{im}}\frac{\exp\left(\delta_{jm}+\mu_{ijm}\right)}{1+\sum_{k=1}^{J}\exp\left(\delta_{km}+\mu_{ikm}\right)}dF\left(\mu_{im}\right)\right) \tag{17}$$

  where S_{jm} is the ovserved market share of product j in market m.
- The above iterative procedure allows us to recover ξ_{jm} and evaluate (and eventually minimize) the moment conditions.

BLP Instruments (Berry, Levinsohn, and Pakes, 1995)

Idea

Use observable characteristics of own and rival products as instruments for price.

- Price is correlated with unobserved quality ξ_{jt}, but also depends on market competition.
- Construct instruments from observed product characteristics:

$$Z_{kjm} = \frac{1}{J-1} \sum_{j'} |X_{kjm} - X_{kj'm}|, \quad \text{When don't have too many parameters}$$

$$Z_{kjj'm} = |X_{kjm} - X_{kj'm}| \quad \text{When need higher number of instrumetns}$$

- Intuition:
  - Rivals' features influence p_{jm} via competitive pressure.
  - Rivals' x_{km} are exogenous to ξ_{jm}.
  - Not exogeneous if attributes are endogeneous

Hausman–Nevo Instruments (1996, 2000)

Idea

Use prices of the same product in other markets as instruments.

- When product j appears in multiple markets t, use:

$$Z_{jm}^{HN} = \frac{1}{T-1} \sum_{s \neq m} p_{js}$$

- Rationale:

$$\begin{aligned} \text{Cov}(p_{jm}, p_{j,-m}) &> 0 \quad \text{(common costs)} \\ \text{Cov}(\xi_{jm}, \xi_{j,-m}) &\approx 0 \quad \text{(local demand shocks)} \end{aligned}$$

- Valid if unobserved demand shocks are market-specific but cost shocks are correlated across markets.

Waldfogel Instruments

Idea

Instrument local price with market-level demographics (or demand shifters) from *other* markets/locations.

- Let $\bar{D}_s$ be a vector of market mean of demographics (income, population, age shares, etc.) in market s.
- For product j in market m construct:

$$Z_{jt}^{W} = \frac{1}{M-1} \sum_{s \neq m} w_{js} \, \bar{D}_s$$

  where w_{js} are optional weights (e.g. 1, distance-based, or market-size weights).
- Intuition:
  - Demographics in other markets predict how firms set prices nationally or across markets (common pricing rules, cost pass-through).
  - Because these are *other* markets, they are plausibly uncorrelated with the local unobserved demand shock ξ_{jm}.
- Main threat: national shocks (advertising, brand reputation) or spatially correlated unobservables that link D_s to ξ_{jm}.