EC708 Discussion 10
Multinomial Choice

Yan Liu

Department of Economics
Boston University

April 14, 2023

Outline

1. Mixed Logit
2. Endogeneity and the BLP Approach

Table of Contents

1 Mixed Logit

2 Endogeneity and the BLP Approach

Mixed Logit

Definition

A mixed multinomial logit model is any model whose choice probabilities can be expressed as

$$P(Y_t = j|X_t) = \int \frac{e^{X'_{tj}\beta}}{\sum_{\ell=1}^{J} e^{X'_{tj}\beta}} f(\beta) d\beta.$$

- Standard logit is special case: $f(\beta) = 1$ for $\beta = b$ and 0 for $\beta \neq b$
- $f(\beta)$ can be discrete $\Rightarrow$ latent class model
- $f(\beta)$ is continuous in most applications. We can allow for
  - correlated normal draws: $\beta \sim N(\bar{\beta}, \Sigma)$
  - non-normal distributions: lognormal, uniform, exponential, etc.

Usually, $f(\beta)$ is parameterized as $f(\beta; \theta)$. We are interested in estimating θ.

Mixed Logit

Random Coefficients Interpretation

Mixed logit can be derived from a random coefficients specification:

$$\pi_{tj} = X_{tj}'\beta_t + U_{tj},$$

where $\beta_t \sim f(\cdot;\theta), U_{tj}, j = 1, \ldots, J \overset{\text{i.i.d.}}{\sim}$ Type I EV, and $\beta_t \perp U_t$.

- If the researcher observed β_t, the choice probability conditional on β_t is standard logit.
- However, the researcher does not know β_t. The unconditional choice probability is

$$P(Y_t = j|X_t) = \int \frac{e^{X_{tj}'\beta}}{\sum_{\ell=1}^{J} e^{X_{t\ell}'\beta}} f(\beta;\theta) d\beta.$$

Mixed Logit

Error Component Interpretation

Utility is specified as

$$\pi_{tj} = X_{tj}'\alpha + \underbrace{Z_{tj}'\mu_t + U_{tj}}_{\tilde{U}_{tj}},$$

where $U_{tj}, j = 1, \ldots, J \overset{\text{i.i.d.}}{\sim}$ Type I EV, μ_t is unobserved, and $E[\mu_t] = 0$.

- For standard logit, Z_{tj} is identically zero.
- With nonzero Z_{tj}, utility is correlated over alternatives:

$$\mathrm{Cov}(\tilde{U}_{tj}, \tilde{U}_{tk}) = E(Z_{tj}'\mu_t + U_{tj})(Z_{tk}'\mu_t + U_{tk}) = Z_{tj}'E[\mu_t\mu_t']Z_{tk}.$$

- Can obtain nested logit with K non-overlapping nests: specify $Z_{tj}'\mu_t = \sum_{k=1}^{K} \mu_{tk}d_{jk}$, where $d_{jk} = 1$ if j is in nest k and 0 otherwise.

Mixed Logit

Substitution Patterns

- At the individual-level decision-making problem, independence of irrelevant alternatives (IIA) is still present.
- At the aggregate level, IIA property disappears: $\frac{P(Y_t=j|X_t)}{P(Y_t=k|X_t)} = \frac{\int (e^{X'_{tj}\beta}/\sum_{\ell=1}^J e^{X'_{t\ell}\beta}) f(\beta;\theta)d\beta}{\int (e^{X'_{tk}\beta}/\sum_{\ell=1}^J e^{X'_{t\ell}\beta}) f(\beta;\theta)d\beta}$ depends on all the data.
- Own-price and cross-price elasticity ($X_{tj}^{(1)}$ is price):

$$\kappa_{jj}(x) = \frac{\partial p_j(x)/\partial x_j^{(1)}}{p_j(x)/x_j^{(1)}} = \frac{x_j^{(1)}}{p_j(x)} \int \beta^{(1)} s_j(x)(1 - s_j(x)) f(\beta;\theta)d\beta,$$

$$\kappa_{kj}(x) = \frac{\partial p_k(x)/\partial x_j^{(1)}}{p_k(x)/x_j^{(1)}} = -\frac{x_j^{(1)}}{p_k(x)} \int \beta^{(1)} s_j(x) s_k(x) f(\beta;\theta)d\beta,$$

where $p_j(x) = P(Y_t = j|X_t = x)$ and $s_j(x) = e^{x'_j\beta}/\sum_{\ell=1}^J e^{x'_\ell\beta}$.

Mixed Logit

Approximation to Any Random Utility Model

McFadden and Train (2000) show that any random utility model (RUM) can be approximated by a mixed logit. Suppose the true RUM is

$$\pi_{tj} = Z'_{tj}\alpha_t, \quad \alpha_t \sim f(\alpha).$$

(We can obtain the traditional notation $\pi_{tj} = X'_{tj}\beta_t + U_{tj}$ by letting $Z_{tj} = (X'_{tj}, e'_j)'$, $\alpha_t = (\beta'_t, U'_t)'$, and $f(\alpha)$ be the joint density of β_t and U_t.) The unconditional choice probability is

$$Q_{tj} = \int 1\{Z'_{tj}\alpha > Z'_{t\ell}\alpha \; \forall j \neq \ell\} f(\alpha) d\alpha.$$

Mixed Logit

Approximation to Any Random Utility Model

We can approximate the true Q_{tj} with a mixed logit.

1. Scale utility by λ so that $\pi^*_{tj} = Z'_{tj}(\alpha_t/\lambda)$.
2. Add an i.i.d. extreme value term ε_{tj}.
3. The mixed logit probability based on this utility is

$$P_{tj} = \int \frac{e^{Z'_{tj}(\alpha/\lambda)}}{\sum_\ell e^{Z'_{t\ell}(\alpha/\lambda)}} f(\alpha) d\alpha.$$

As $\lambda \to 0$, α_t/λ grows large, and $\frac{e^{Z'_{tj}(\alpha/\lambda)}}{\sum_\ell e^{Z'_{t\ell}(\alpha/\lambda)}}$ approaches $1\{Z'_{tj}\alpha > Z'_{t\ell}\alpha \; \forall j \neq \ell\}$, i.e. P_{tj} approaches Q_{tj}.

Simulated Maximum Likelihood (SML) estimator $\hat{\theta}_{SML}$ maximizes

$$\bar{L}_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{J} 1\{Y_t = j\} \ln \left(\frac{1}{S} \sum_{s=1}^{S} \frac{e^{X_{tj}' \beta^s}}{\sum_{\ell=1}^{J} e^{X_{t\ell}' \beta^s}} \right),$$

where $\beta^s, s = 1, \ldots, S \overset{\text{i.i.d.}}{\sim} f(\cdot; \theta)$.

Remark 1. $\hat{\theta}_{SML}$ is consistent if $T, S \to \infty$.

1. As $S \to \infty$, $\frac{1}{S} \sum_{s=1}^{S} \frac{e^{X_{tj}' \beta^s}}{\sum_{\ell=1}^{J} e^{X_{t\ell}' \beta^s}} \xrightarrow{p} E_{\beta^s} \left[\frac{e^{X_{tj}' \beta^s}}{\sum_{\ell=1}^{J} e^{X_{t\ell}' \beta^s}} \right] = P(Y_t = j | X_t)$
2. Hence, $\text{plim}_{T,S \to \infty} \bar{L}_T(\theta) = E\big[\sum_{j=1}^{J} 1\{Y_t = j\} \ln P(Y_t = j | X_t) \big| X_t \big]$, whose maximizer is θ_0.

Mixed Logit

Estimation: SML

Remark 2. $\hat{\theta}_{SML}$ is inconsistent if S is fixed and $T \to \infty$.

- As $T \to \infty$, the probability limit of $\bar{L}_T(\theta)$ is

$$E\Big[\sum_{j=1}^{J} 1\{Y_t = j\} E_{\beta^1,\dots,\beta^S}\Big[\ln\Big(\frac{1}{S}\sum_{s=1}^{S}\frac{e^{X_{tj}'\beta^s}}{\sum_{\ell=1}^{J} e^{X_{t\ell}'\beta^s}}\Big)\Big]\Big|X_t\Big].$$

- Log and integral do not commute:

$$\begin{aligned} E_{\beta^1,\dots,\beta^S}\Big[\ln\Big(\frac{1}{S}\sum_{s=1}^{S}\frac{e^{X_{tj}'\beta^s}}{\sum_{\ell=1}^{J} e^{X_{t\ell}'\beta^s}}\Big)\Big] &\neq \ln E_{\beta^1,\dots,\beta^S}\Big[\frac{1}{S}\sum_{s=1}^{S}\frac{e^{X_{tj}'\beta^s}}{\sum_{\ell=1}^{J} e^{X_{t\ell}'\beta^s}}\Big] \\ &= \ln P(Y_t = j|X_t). \end{aligned}$$

Table of Contents

1 Mixed Logit

2 Endogeneity and the BLP Approach

Discrete-choice demand model at the individual household level:

$$\pi_{tj} = \underbrace{X_j'\beta - \alpha p_j + \xi_j}_{\equiv \delta_j \text{ "mean utility"}} + U_{tj}$$

- i indexes household, j indexes brand
- Both ξ_j and U_{tj} are unobserved
- $U_{tj}, j = 1, \ldots, J \overset{\text{i.i.d.}}{\sim}$ Type I EV
- ξ_j is "unobserved quality" and correlated with $p_j \Rightarrow$ endogeneity
- Aggregate data: contains X_j, p_j and market shares $\hat{s}_j$ across j.

Berry (1994) suggests an IV-based estimation approach.

- Assume there exist IVs Z_j such that $E[\xi_j Z_j] = 0$.
- Sample moment condition

$$\frac{1}{J}\sum_{j=1}^{J} \xi_j Z_j = \frac{1}{J}\sum_{j=1}^{J} (\delta_j - X_j' \beta + \alpha p_j) Z_j$$

  which converges (as $J \to \infty$) to zero at true α_0, β_0.

However, we do not know δ_j. **Solution:** a two-step approach

BLP

Berry (1994)

First step: Inversion

Equate observed market shares $\hat{s}_j$ to predicted market shares

$$\tilde{s}_j(\delta_0, \delta_1, \ldots, \delta_J) = \frac{e^{\delta_j}}{1 + \sum_{\ell=1}^{J} e^{\delta_\ell}}, \text{ where } \delta_0 = 0.$$

$\Rightarrow$ Invert a system of J nonlinear equations:

$$\begin{aligned}
\hat{s}_0 &= \tilde{s}_0(0, \hat{\delta}_1, \ldots, \hat{\delta}_J) \\
\hat{s}_1 &= \tilde{s}_1(0, \hat{\delta}_1, \ldots, \hat{\delta}_J) \Rightarrow \hat{\delta}_1 = \ln \hat{s}_1 - \ln \hat{s}_0 \\
&\vdots \\
\hat{s}_J &= \tilde{s}_J(0, \hat{\delta}_1, \ldots, \hat{\delta}_J) \Rightarrow \hat{\delta}_J = \ln \hat{s}_J - \ln \hat{s}_0
\end{aligned}$$

BLP

Berry (1994)

Second step: IV estimation

Use estimated $\hat{\delta}_j$'s to calculate sample moment condition:

$$\frac{1}{J}\sum_{j=1}^{J}\xi_j Z_j = \frac{1}{J}\sum_{j=1}^{J}(\hat{\delta}_j - X_j'\beta + \alpha p_j)Z_j.$$

GMM can be implemented by running IV regression of

$$\ln \hat{s}_j - \ln \hat{s}_0 = X_j'\beta - \alpha p_j + \xi_j \quad (\text{"logistic IV regression"}).$$

BLP

Instruments

Berry, Levinsohn, and Pakes (1995) propose to use as Z_j characteristics of rival products. **Intuition:**

- Oligopolistic competition makes firm j set price p_j as a function of characteristics of products by firms $k \neq j$.
- Characteristics of rival products should not affect households' valuation of brand j.

BLP Approach

Random-Coefficients Logit Model

Assume that utility function is

$$\pi_{tj} = X_j' \beta_t - \alpha_t p_j + \xi_j + U_{tj},$$

Assume $(\alpha_t, \beta_t) \perp U_t$ and $(\alpha_t, \beta_t')' \sim N((\bar{\alpha}, \bar{\beta}')', \Sigma)$. Decompose

$$\pi_{tj} = \underbrace{X_j' \bar{\beta} - \bar{\alpha} p_j + \xi_j}_{=\delta_j} + X_j'(\beta_t - \bar{\beta}) - (\alpha_t - \bar{\alpha}) p_j + U_{tj}.$$

The simple multinomial logit inversion method will not work.

$\Rightarrow$ Berry et al. (1995) propose a "nested" estimation algorithm.

BLP Approach

Inner Loop: The Contraction

- Predicted market share:

$$\tilde{s}_j^{\mathrm{RC}}(\delta_1, \ldots, \delta_J; \bar{\alpha}, \bar{\beta}, \Sigma) \equiv \int\int \frac{e^{\delta_j + X_j'(\beta-\bar{\beta})-(\alpha-\bar{\alpha})p_j}}{1+\sum_{\ell=1}^{J} e^{\delta_\ell + X_\ell'(\beta-\bar{\beta})-(\alpha-\bar{\alpha})p_\ell}} dF(\alpha, \beta).$$

- At each trial value of $(\bar{\alpha}, \bar{\beta}, \Sigma)$, adjust δ_j iteratively by

$$\delta_j^{k+1} = \delta_j^k + \ln \hat{s}_j - \ln \tilde{s}_j^{\mathrm{RC}}(\delta_1^k, \ldots, \delta_J^k; \bar{\alpha}, \bar{\beta}, \Sigma).$$

  - Berry et al. (1995) showed that the iterative adjustment process is a contraction that guarantees convergence.

- Output: $\hat{\delta}_j(\bar{\alpha}, \bar{\beta}, \Sigma), j = 1, \ldots, J$

BLP Approach

Inner Loop: The Contraction

Berry et al. (1995) propose simulation methods to compute the multidimensional integral:

$$\tilde{s}_j^{\mathrm{RC}}(\delta_1, \ldots, \delta_J; \bar{\alpha}, \bar{\beta}, \Sigma) \approx \frac{1}{S}\sum_{s=1}^{S} \frac{e^{\delta_j + X_j'(\beta^s - \bar{\beta}) - (\alpha^s - \bar{\alpha})p_j}}{1 + \sum_{\ell=1}^{J} e^{\delta_\ell + X_\ell'(\beta^s - \bar{\beta}) - (\alpha^s - \bar{\alpha})p_\ell}},$$

where $(\alpha^s, \beta^{s\prime})' \overset{\text{i.i.d.}}{\sim} N((\bar{\alpha}, \bar{\beta}')', \Sigma)$.

BLP Approach

Outer Loop: Estimation by GMM

Construct the GMM objective

$$\Big(\frac{1}{J}\sum_{j=1}^{J}\hat{\xi}_j(\bar{\alpha},\bar{\beta},\Sigma)Z_j\Big)'W_J\Big(\frac{1}{J}\sum_{j=1}^{J}\hat{\xi}_j(\bar{\alpha},\bar{\beta},\Sigma)Z_j\Big),$$

where $\hat{\xi}_j(\bar{\alpha},\bar{\beta},\Sigma)=\hat{\delta}_j(\bar{\alpha},\bar{\beta},\Sigma)-X_j'\bar{\beta}+\bar{\alpha}p_j$.

- Necessary condition for identification:

$$\dim(Z_j)\geq\dim(\bar{\alpha},\bar{\beta},\Sigma)>\dim(\bar{\alpha},\bar{\beta})=\dim(X_j,p_j).$$

  - Even without price endogeneity, still need additional IVs to identify Σ.
- Can concentrate out $(\bar{\alpha},\bar{\beta})$ and search over Σ.