Data-Driven Nests in Discrete Choice Models

Midwest IO Fest

September, 2024

Milena Almagro
(Chicago Booth & NBER)

Kenneth Lai
(NYU)

Elena Manresa
(NYU)

Introduction

Modelling product demand

Discrete choice models are the workhorse in demand estimation with random utility

- Utility is driven by observables + unobservable idiosyncratic taste shock, typically i.i.d.
- Agents choose the alternative with highest utility

Modelling product demand

Discrete choice models are the workhorse in demand estimation with random utility

- Utility is driven by observables + unobservable idiosyncratic taste shock, typically i.i.d.
- Agents choose the alternative with highest utility

Examples:

- **Industrial Organization**: market shares and Logit shocks
- **Trade**: expenditure shares and Fréchet shocks

Modelling product demand

Discrete choice models are the workhorse in demand estimation with random utility

- Utility is driven by observables + unobservable idiosyncratic taste shock, typically i.i.d.
- Agents choose the alternative with highest utility

Examples:

- **Industrial Organization**: market shares and Logit shocks
- **Trade**: expenditure shares and Fréchet shocks

Pros:

- Closed form solutions of choice probability + low number of parameters

Modelling product demand

Discrete choice models are the workhorse in demand estimation with random utility

- Utility is driven by observables + unobservable idiosyncratic taste shock, typically i.i.d.
- Agents choose the alternative with highest utility

Examples:

- **Industrial Organization**: market shares and Logit shocks
- **Trade**: expenditure shares and Fréchet shocks

Pros:

- Closed form solutions of choice probability + low number of parameters

Cons:

- Restricted substitution patterns + requires independence assumptions across products

Toward more flexible substitution patterns

Some proposed alternatives

– Random Coefficients: Logit with heterogeneity in preferences across consumers

  + Flexible substitution patterns
  - Non-linear estimation: numerical integration, no closed-form demand, numerical instability
    Dube, Fox and Su (2012), Knittel and Metaxoglou (2014)
  - Distributional assumptions on heterogeneity

– Nested Structures: Natural extension of i.i.d. shocks

  + Closed form solutions for choice probability
  + Suited to capture market segmentation
  - Nests need to be specified ex-ante
  - Still somewhat restrictive substitution patterns

This paper: Estimating nests in discrete choice models

- Methodology to estimate the nest structure as well as preference parameters
- Nest structure is recovered from **aggregate share data**
- Two-step estimation procedure:
  1. Use k-means clustering to estimate the nest structure
  2. Estimate model parameters as if the groups where known
- We exploit the structure of the model, the availability of many markets and of many products

Related Literature

- **Discrete Choice Models of Random Utility with Nests:** McFadden (1978, 1981), Berry (1994), Verboven (1996), Cardell (1997), McFadden and Train (2000), Grigolon and Verboven (2014), Fosgerau, Monardo, and De Palma (2022)

- **Empirical Applications of Nesting Structures:**
  - Industrial Organization: Einav (2007), Grennan (2013), Ciliberto & Williams (2014), Conlon & Rao (2016), Miller & Weinberg (2017)...
  - Trade/Spatial/Urban: Goldberg (1995), Broda and Weinstein (2006), Atkenson and Burnstein (2008), Couture, Gaubert, Handbury and Hurst (2023), Bordeu (2024)...

- **Group Fixed Effect Estimator:** Han & Moon (2010), Bonhomme & Manresa (2015)

- **Alternative Grouping Structure:** Fosgerau, Monardo, & De Palma (2022), Hortacsu, Lieber, Monardo & de Paula (ongoing)

Outline

Empirical model

Discrete choice model with nested Logit shocks

- Consider the indirect utility model for agent i when choosing j:

$$V_{ij} = \delta_j + \varepsilon_{ij}$$

- Choice of j based on the maximization of the utiliy:

$$\mathbb{P}_j = \mathbb{P}(V_{ij} > V_{ij'} \quad \forall j' \neq j)$$

Discrete choice model with nested Logit shocks

- Consider the indirect utility model for agent i when choosing j:

$$V_{ij} = \delta_j + \varepsilon_{ij}$$

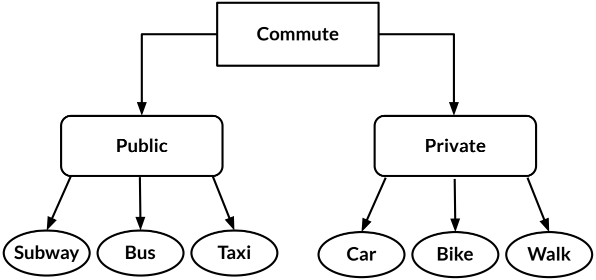
- Choice of j based on the maximization of the utiliy:

$$\mathbb{P}_j = \mathbb{P}(V_{ij} > V_{ij'} \quad \forall j' \neq j)$$

- Assume products are partitioned in K groups, and $(\varepsilon_{i1}, \ldots, \varepsilon_{iJ}) \sim \exp\Big(-\sum_{k=1}^{K} (\sum_{j \in B_K} e^{-\frac{\epsilon_j}{\sigma^{k(j)}}})^{\sigma^{k(j)}} \Big)$:

$$\mathbb{P}_j = \frac{e^{\frac{\delta_j}{\sigma^{k(j)}}} \Big(\sum_{d \in B_{k(j)}} e^{\frac{\delta_d}{\sigma^{k(j)}}} \Big)^{\sigma^{k(j)}-1}}{\sum_{l=1}^{K} \Big(\sum_{d \in B_l} e^{\frac{\delta_d}{\sigma^l}} \Big)^{\sigma^l}}. \tag{1}$$

Nested Logit as sequential choice



Choice of option j within nest $k(j)$

$$\mathbb{P}_j = \underbrace{\left(\frac{e^{\frac{\delta_j}{\sigma^{k(j)}}}}{\sum_{d \in B_{k(j)}} e^{\frac{\delta_d}{\sigma^{k(j)}}}} \right)}_{\mathbb{P}_{j|k(j)}} \underbrace{\left(\frac{\left(\sum_{d \in B_{k(j)}} e^{\frac{\delta_d}{\sigma^{k(j)}}} \right)^{\sigma^{k(j)}}}{\sum_{l=1}^{K} \left(\sum_{d \in B_l} e^{\frac{\delta_d}{\sigma^l}} \right)^{\sigma^l}} \right)}_{\mathbb{P}_{k(j)}}$$

Nested Logit and substitution patterns

Correlation across products within nest:

$$V_{ij} = \delta_j + \varsigma^i_{k(j)} + \epsilon_{ij}$$

- Conditional on nest, shocks are still i.i.d.

Nested Logit and substitution patterns

Correlation across products within nest:

$$V_{ij} = \delta_j + \varsigma^i_{k(j)} + \epsilon_{ij}$$

- Conditional on nest, shocks are still i.i.d.

Group correlation allows for more flexible substitution patterns

$$\mathcal{E}^{j'}_j = \begin{cases} -\beta_{j'}\mathbb{P}_{j'}p_{j'} & \text{if } j' \notin B_{k(j)} \\ (\sigma_k - 1)\frac{\beta_p}{\sigma_k}\mathbb{P}_{j'|k}p_{j'} - \beta_p\mathbb{P}_{j'}p_{j'} & \text{if } j' \in B_{k(j)}. \end{cases}$$

Some notation

Define $\mathrm{IV}^k \equiv \sum_{d \in B_k} e^{\frac{\delta_d}{\sigma^k}}$ and $\mathrm{IV} \equiv \sum_{l=1}^{K} \left(\sum_{d \in B_l} e^{\frac{\delta_d}{\sigma^l}} \right)^{\sigma^l}$

Some notation

Define $\text{IV}^k \equiv \sum_{d \in B_k} e^{\frac{\delta_d}{\sigma^k}}$ and $\text{IV} \equiv \sum_{l=1}^{K} \left(\sum_{d \in B_l} e^{\frac{\delta_d}{\sigma^l}} \right)^{\sigma^l}$

Then:

$$\mathbb{P}_j = \left(\frac{e^{\frac{\delta_j}{\sigma^{k(j)}}}}{\sum_{d \in B_{k(j)}} e^{\frac{\delta_d}{\sigma^{k(j)}}}} \right) \left(\frac{\left(\sum_{d \in B_{k(j)}} e^{\frac{\delta_d}{\sigma^{k(j)}}} \right)^{\sigma^{k(j)}}}{\sum_{l=1}^{K} \left(\sum_{d \in B_l} e^{\frac{\delta_d}{\sigma^l}} \right)^{\sigma^l}} \right) = \frac{e^{\frac{\delta_j}{\sigma^{k(j)}}} (IV^k)^{\sigma^{k(j)}-1}}{IV}$$

Some notation

Define $\text{IV}^k \equiv \sum_{d\in B_k} e^{\frac{\delta_d}{\sigma^k}}$ and $\text{IV} \equiv \sum_{l=1}^{K} \big(\sum_{d\in B_l} e^{\frac{\delta_d}{\sigma^l}}\big)^{\sigma^l}$

Then:

$$\mathbb{P}_j = \left(\frac{e^{\frac{\delta_j}{\sigma^{k(j)}}}}{\sum_{d\in B_{k(j)}} e^{\frac{\delta_d}{\sigma^{k(j)}}}}\right)\left(\frac{\big(\sum_{d\in B_{k(j)}} e^{\frac{\delta_d}{\sigma^{k(j)}}}\big)^{\sigma^{k(j)}}}{\sum_{l=1}^{K}\big(\sum_{d\in B_l} e^{\frac{\delta_d}{\sigma^l}}\big)^{\sigma^l}}\right) = \frac{e^{\frac{\delta_j}{\sigma^{k(j)}}}(IV^k)^{\sigma^{k(j)}-1}}{IV}$$

Taking logs:

$$\log \mathbb{P}_j = \frac{\delta_j}{\sigma^{k(j)}} + (\sigma^{k(j)} - 1)\log \text{IV}^{k(j)} - \log \text{IV}$$

Outline

Intuition of nest identification

For simplicity, assume $\delta_j = \beta x_j$, $k(0) = \{0\}$ and $\delta_0 = 0$.

Intuition of nest identification

For simplicity, assume $\delta_j = \beta x_j$, $k(0) = \{0\}$ and $\delta_0 = 0$.

Denote $\theta_{k(j)} \equiv \frac{\beta}{\sigma^{k(j)}}$ and $\lambda_{k(j)} \equiv (\sigma^{k(j)} - 1) \log \mathrm{IV}^{k(j)}$.

Intuition of nest identification

For simplicity, assume $\delta_j = \beta x_j$, $k(0) = \{0\}$ and $\delta_0 = 0$.

Denote $\theta_{k(j)} \equiv \frac{\beta}{\sigma^{k(j)}}$ and $\lambda_{k(j)} \equiv (\sigma^{k(j)} - 1) \log \mathrm{IV}^{k(j)}$.

It follows:

$$\begin{aligned}\log \mathbb{P}_j - \log \mathbb{P}_0 &= \frac{\beta x_j}{\sigma^{k(j)}} + (\sigma^{k(j)} - 1) \log \mathrm{IV}^{k(j)} \\ &= \theta_{k(j)} x_j + \lambda_{k(j)},\end{aligned}$$

Intuition of nest identification

For simplicity, assume $\delta_j = \beta x_j$, $k(0) = \{0\}$ and $\delta_0 = 0$.

Denote $\theta_{k(j)} \equiv \frac{\beta}{\sigma^{k(j)}}$ and $\lambda_{k(j)} \equiv (\sigma^{k(j)} - 1) \log \text{IV}^{k(j)}$.

It follows:

$$
\begin{aligned}
\log \mathbb{P}_j - \log \mathbb{P}_0 &= \frac{\beta x_j}{\sigma^{k(j)}} + (\sigma^{k(j)} - 1) \log \text{IV}^{k(j)} \\
&= \theta_{k(j)} x_j + \lambda_{k(j)},
\end{aligned}
$$

$\Rightarrow$ **Key observation:** Group-specific intercept and slope common for products in the same group!

Intuition of nest identification

For simplicity, assume $\delta_j = \beta x_j$, $k(0) = \{0\}$ and $\delta_0 = 0$.

Denote $\theta_{k(j)} \equiv \frac{\beta}{\sigma^{k(j)}}$ and $\lambda_{k(j)} \equiv (\sigma^{k(j)} - 1) \log \text{IV}^{k(j)}$.

It follows:

$$\begin{aligned} \log \mathbb{P}_j - \log \mathbb{P}_0 &= \frac{\beta x_j}{\sigma^{k(j)}} + (\sigma^{k(j)} - 1) \log \text{IV}^{k(j)} \\ &= \theta_{k(j)} x_j + \lambda_{k(j)}, \end{aligned}$$

$\Rightarrow$ **Key observation:** Group-specific intercept and slope common for products in the same group!

$\Rightarrow$ **Intuition:** The marginal effect of extra x_j varies by nest

Introducing markets, error, and prices

Let $j = 0, 1, \ldots, J$ denote products, $m = 1, \ldots, M$ markets and p_{jm} prices, so that:

$$\delta_{jm} = \beta_p p_{jm} + \beta_x x_{jm} + \xi_{jm},$$

Introducing markets, error, and prices

Let $j = 0, 1, \ldots, J$ denote products, $m = 1, \ldots, M$ markets and p_{jm} prices, so that:

$$\delta_{jm} = \beta_p p_{jm} + \beta_x x_{jm} + \xi_{jm},$$

Our estimation equation becomes:

$$\log \mathbb{P}_{jm} - \log \mathbb{P}_{0m} = \beta^{k(j)} p_{jm} + \beta_x^{k(j)} x_{jm} + \lambda_m^{k(j)} + \tilde{\xi}_{jm},$$

where $\beta_p^{k(j)} \equiv \frac{\beta^p}{\sigma^{k(j)}}, \beta_x^{k(j)} \equiv \frac{\beta^x}{\sigma^{k(j)}}, \lambda_m^{k(j)} \equiv (\sigma^{k(j)} - 1) \log \mathrm{IV}_m^{k(j)}$, and $\tilde{\xi}_j \equiv \frac{\xi_j}{\sigma^{k(j)}}$

Introducing markets, error, and prices

Let $j = 0, 1, \ldots, J$ denote products, $m = 1, \ldots, M$ markets and p_{jm} prices, so that:

$$\delta_{jm} = \beta_p p_{jm} + \beta_x x_{jm} + \xi_{jm},$$

Our estimation equation becomes:

$$\log \mathbb{P}_{jm} - \log \mathbb{P}_{0m} = \beta^{k(j)} p_{jm} + \beta_x^{k(j)} x_{jm} + \lambda_m^{k(j)} + \tilde{\xi}_{jm},$$

where $\beta_p^{k(j)} \equiv \frac{\beta^p}{\sigma^{k(j)}}, \beta_x^{k(j)} \equiv \frac{\beta^x}{\sigma^{k(j)}}, \lambda_m^{k(j)} \equiv (\sigma^{k(j)} - 1) \log \mathrm{IV}_m^{k(j)}$, and $\tilde{\xi}_j \equiv \frac{\xi_j}{\sigma^{k(j)}}$

For now, let's assume $\mathbb{E}[\xi_{jm} p_{jm}] = 0$
$\rightarrow$ Don't worry, we will relax this later...

Outline

Estimation of Groups

Group-fixed effect estimator defined by the following clustering problem:

$$\underset{\substack{k(1),\ldots,k(J)\\ \beta^1,\ldots,\beta^K,\lambda_1^1,\ldots,\lambda_M^K}}{\arg\min} \sum_{m=1}^{M}\sum_{j=1}^{J}\left(\log\left(\frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}}\right) - (\beta_p^k p_{jm} + \beta_x^k x_{jm} + \lambda_m^k)\right)^2$$

Combinatorial, non-convex problem!

Estimation of Groups

Group-fixed effect estimator defined by the following clustering problem:

$$\underset{\substack{k(1),\ldots,k(J)\\ \beta^1,\ldots,\beta^K,\lambda_1^1,\ldots,\lambda_M^K}}{\arg\min} \sum_{m=1}^{M}\sum_{j=1}^{J}\left(\log\left(\frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}}\right) - (\beta_p^k p_{jm} + \beta_x^k x_{jm} + \lambda_m^k)\right)^2$$

Combinatorial, non-convex problem!

Solution: two-step algorithm

1. Classify products using clustering algorithm following Bonhomme and Manresa (2015)
2. Conditional on classification, estimate preference parameters β and σ following Berry (1994)

Why two steps?

Two-step strategy

First Step: Classification (Bonhomme and Manresa, 2015)

1. Let $(\beta^{1,0}, \dots, \beta^{K,0}, \lambda_1^{K,0}, \dots, \lambda_M^{K,0})$ be a starting value.

Two-step strategy

First Step: Classification (Bonhomme and Manresa, 2015)

1. Let $(\beta^{1,0}, \dots, \beta^{K,0}, \lambda_1^{K,0}, \dots, \lambda_M^{K,0})$ be a starting value.
2. For $(\beta^{1,s}, \dots, \beta^{K,s}, \lambda_1^{K,s}, \dots, \lambda_M^{K,s})$, compute for all $j \in J$:

$$k(j)^{s+1} = \underset{k \in \{1,\dots,K\}}{\arg\min} \sum_{m=1}^{M} \Big(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_p^{k,s} p_{jm} + \beta_x^{k,s} x_{jm} + \lambda_m^{k,s}) \Big)^2,$$

to compute grouping structure $\mathcal{B}^{s+1}$.

Two-step strategy

First Step: Classification (Bonhomme and Manresa, 2015)

1. Let $(\beta^{1,0}, \dots, \beta^{K,0}, \lambda_1^{K,0}, \dots, \lambda_M^{K,0})$ be a starting value.
2. For $(\beta^{1,s}, \dots, \beta^{K,s}, \lambda_1^{K,s}, \dots, \lambda_M^{K,s})$, compute for all $j \in J$:

$$k(j)^{s+1} = \underset{k\in\{1,\dots,K\}}{\arg\min} \sum_{m=1}^{M} \Big(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_p^{k,s} p_{jm} + \beta_x^{k,s} x_{jm} + \lambda_m^{k,s}) \Big)^2,$$

   to compute grouping structure $\mathcal{B}^{s+1}$.

3. Compute:

$$(\beta^{1,s+1}, \dots, \beta^{K,s+1}, \lambda_1^{K,s+1}, \dots, \lambda_M^{K,s+1}) =$$

$$\underset{\beta^1,\dots,\beta^K,\lambda_1^1,\dots,\lambda_M^K}{\arg\min} \sum_{j=1}^{J} \sum_{m=1}^{M} \Big(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_p^{k(j),s+1} p_{jm} + \beta_x^{k(j),s+1} x_{jm} + \lambda_m^{k(j),s+1}) \Big)^2$$

Two-step strategy

First Step: Classification (Bonhomme and Manresa, 2015)

1. Let $(\beta^{1,0},\ldots,\beta^{K,0},\lambda_1^{K,0},\ldots,\lambda_M^{K,0})$ be a starting value.
2. For $(\beta^{1,s},\ldots,\beta^{K,s},\lambda_1^{K,s},\ldots,\lambda_M^{K,s})$, compute for all $j \in J$:

$$k(j)^{s+1} = \underset{k\in\{1,\ldots,K\}}{\arg\min} \sum_{m=1}^{M} \left(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_p^{k,s} p_{jm} + \beta_x^{k,s} x_{jm} + \lambda_m^{k,s}) \right)^2,$$

   to compute grouping structure $\mathcal{B}^{s+1}$.

3. Compute:

$$(\beta^{1,s+1},\ldots,\beta^{K,s+1},\lambda_1^{K,s+1},\ldots,\lambda_M^{K,s+1}) =$$

$$\underset{\beta^1,\ldots,\beta^K,\lambda_1^1,\ldots,\lambda_M^K}{\arg\min} \sum_{j=1}^{J} \sum_{m=1}^{M} \left(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_p^{k(j),s+1} p_{jm} + \beta_x^{k(j),s+1} x_{jm} + \lambda_m^{k(j),s+1}) \right)^2$$

4. Repeat until convergence of parameters.

Two-step strategy

Second step estimation

Once we have grouping structure $\{k(1), ..., k(J)\}$, estimate β and σ as if those were known

Two-step strategy

Second step estimation

Once we have grouping structure $\{k(1), \ldots, k(J)\}$, estimate β and σ as if those were known

We follow Berry (1994):

$$\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} = \beta_p p_{jm} + \beta_x x_{jm} + (\sigma^{k(j)} - 1) \log \mathbb{P}_{j,m|k(j)} + \xi_{jm}$$

Observation: Need to instrument for $\log \mathbb{P}_{j,m|k(j)}$, where

$$\mathbb{P}_{j,m|k(j)} = \frac{e^{\frac{\beta_p p_{jm} + \beta_x x_{jm}}{\sigma^{k(j)}}}}{\sum_{d \in B_{k(j)}} e^{\frac{\beta_p p_{jm} + \beta_x x_{jm}}{\sigma^{k(j)}}}}$$

Two-step strategy

Second step estimation

Once we have grouping structure $\{k(1), \ldots, k(J)\}$, estimate β and σ as if those were known

We follow Berry (1994):

$$\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} = \beta_p p_{jm} + \beta_x x_{jm} + (\sigma^{k(j)} - 1) \log \mathbb{P}_{j,m|k(j)} + \xi_{jm}$$

Observation: Need to instrument for $\log \mathbb{P}_{j,m|k(j)}$, where

$$\mathbb{P}_{j,m|k(j)} = \frac{e^{\frac{\beta_p p_{jm} + \beta_x x_{jm}}{\sigma^{k(j)}}}}{\sum_{d \in B_{k(j)}} e^{\frac{\beta_p p_{jm} + \beta_x x_{jm}}{\sigma^{k(j)}}}} \implies z_{jm} = -\log\Big(\sum_{\substack{d \neq j \\ d \in B_{k(j)}}} e^{x_{dm}} \Big)$$

Statistical Properties

Asymptotic Experiment and Conditions

when Nests are Unknown

1. Let $J \to \infty$, $M \to \infty$, and K fixed.
2. We abstract from the noise in estimation in shares coming from a finite population of consumers
3. We consider balanced nests: $J_k = O_p(J)$ for $k = 1, \ldots K$, where $J_k = |B_k|$
4. We consider a sequence $(\sigma_{01,J}, \ldots, \sigma_{0K,J})_{J=1}^{\infty}$ such that:
   (i) $\sigma_{0k,J} \in (0, 1)$,
   (ii) $\sigma_{0\ell,J} - \sigma_{0k,J} = \frac{c_{\ell k,J}}{\log J}$, where $c_{\ell k,J} \to c_{\ell k} \in \mathbb{R}$ as $J \to \infty$,
   (iii) $\sigma_{0\ell,J} \to \sigma_{0\ell}$ where $0 < \sigma_{0\ell} < 1$ for all ℓ.
5. + Additional regularity conditions

Asymptotic Experiment and Conditions

when Nests are Unknown

1. Let $J \to \infty$, $M \to \infty$, and K fixed.
2. We abstract from the noise in estimation in shares coming from a finite population of consumers
3. We consider balanced nests: $J_k = O_p(J)$ for $k = 1, \dots K$, where $J_k = |B_k|$
4. We consider a sequence $(\sigma_{01,J}, \dots, \sigma_{0K,J})_{J=1}^{\infty}$ such that:
   (i) $\sigma_{0k,J} \in (0, 1)$,
   (ii) $\sigma_{0\ell,J} - \sigma_{0k,J} = \frac{c_{\ell k,J}}{\log J}$, where $c_{\ell k,J} \to c_{\ell k} \in \mathbb{R}$ as $J \to \infty$,
   (iii) $\sigma_{0\ell,J} \to \sigma_{0\ell}$ where $0 < \sigma_{0\ell} < 1$ for all ℓ.
5. + Additional regularity conditions

$\Rightarrow$ Asymptotic equivalence with an estimator of the same model where the groups are known

Intuition 1: Normalization and Compact Parameter Space

We re-write our model and multiply both sides of the equation by $\log J$ to obtain:

$$\log JP_{jm} = \frac{\delta_{jm}}{\sigma_{0\ell}} + \underbrace{\log \overline{P}_{k_{0j}m} - \log\left[\frac{1}{J}\sum_{j' \in B_{0k_{0j}}} \exp\left(\frac{\delta_{j'm}}{\sigma_0 k_{0j}}\right)\right]}_{\bar{\zeta}_{0k_{0j}m,J}}$$

where

$$\overline{P}_{\ell m} := \frac{\left(\sum_{j' \in B_{0\ell}} \exp\left(\frac{\delta_{j'm}}{\sigma_{0\ell}}\right)\right)^{\sigma_{0\ell}}}{\sum_{k \in K}\left(\sum_{j' \in B_{0k}} \exp\left(\frac{\delta_{j'm}}{\sigma_{0k}}\right)\right)^{\sigma_{0k}}}.$$

The conditions ensure that $\bar{\zeta}_{0k_{0j}m,J} \to \zeta_{0k_{0j}m}$ in probability uniformly in m, where $\zeta_{0k_{0j}m}$ is finite.

Intuition 2: Misclassification probability

Simplified Example

- Consider the following simplified model with $G = 2$:

$$y_{im} = \alpha^*_{k_i^*} + v_{im}, \qquad k_i \in \{1, 2\}.$$

- We characterize the misclassification probability:

$$\Pr\left(\widehat{k_i}(\alpha) = 2 \middle| k_i^* = 1\right) = \Pr\left((\overline{y}_i - \alpha_2)^2 < (\overline{y}_i - \alpha_1)^2 \middle| k_i^* = 1\right).$$

- If v_{im} are iid normal $(0, \sigma^2)$ and $\alpha_1 < \alpha_2$ then this is:

$$\Pr\left(\overline{v}_i > \frac{\alpha_1 + \alpha_2}{2} - \alpha_1^*\right) \quad = \quad 1 - \Phi\left(\frac{\sqrt{M}}{\sigma}\left(\frac{\alpha_1 + \alpha_2}{2} - \alpha_1^*\right)\right),$$

which vanishes exponentially fast as M increases.

Monte Carlo

Monte Carlo Design: Data

Indirect utility δ_{jm} is given by

$$\delta_{jm} = \beta_p p_{jm} + \beta_x x_{jm} + \xi_{jm},$$

where

- $K = 3$ with $\sigma = (0.2, 0.3, 0.6)$
- Classify products randonmly $k(j) \sim \mathcal{U}\{1, 2, 3\}$
- $\beta_p = -1, \beta_x = 1$
- $\begin{bmatrix} \mu_p^k \\ \mu_x^k \end{bmatrix} \overset{i.i.d.}{\sim} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix} \right)$ across k.
- $\begin{bmatrix} p_{jm} \\ x_{jm} \\ \xi_{jm} \end{bmatrix} \overset{i.i.d.}{\sim} \mathcal{N}\left(\begin{bmatrix} \mu_p^k \\ \mu_x^k \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \right)$ across j and m $\implies$ ξ_{jm} explains 20% of variation in δ_{jm}

Results

Results of 50 Bootstrap iterations

J	M	Runtime	Matched	True	β_p	β_c	σ_1	σ_2	σ_3
				True	-1	1	0.2	0.3	0.6
100	10	00:02	0.996	Mean β	-0.992	0.991	0.189	0.297	0.602
				Std β	0.032	0.033	0.034	0.024	0.007
100	50	00:25	1.0	Mean β	-1.001	0.998	0.2	0.3	0.6
				Std β	0.01	0.011	0.001	0.002	0.003
100	100	01:07	1.0	Mean β	-1.0	1.0	0.2	0.3	0.6
				Std β	0.006	0.007	0.001	0.001	0.003
500	10	00:06	0.995	Mean β	-1.0	0.998	0.199	0.298	0.6
				Std β	0.015	0.02	0.006	0.016	0.003
500	50	07:14	1.0	Mean β	-1.0	0.999	0.2	0.3	0.6
				Std β	0.004	0.004	0.0	0.001	0.001
500	100	29:24	1.0	Mean β	-1.0	0.999	0.2	0.3	0.6
				Std β	0.003	0.003	0.0	0.0	0.001
1000	10	00:12	1.0	Mean β	-1.0	0.999	0.199	0.3	0.6
				Std β	0.007	0.007	0.003	0.001	0.002
1000	50	44:57	1.0	Mean β	-1.0	1.0	0.2	0.3	0.6
				Std β	0.002	0.003	0.0	0.0	0.001
1000	100	11:14	1.0	Mean β	-1.0	0.999	0.2	0.3	0.6
				Std β	0.002	0.002	0.0	0.0	0.001

Extensions

Outline

Choosing K: Cross-validation with Elbow method

So far we have assumed the number of groups is known.

In practice, we can also estimate the number of groups using a N**-fold cross-validation** procedure.

For $k \in \{1, \ldots, K\}$:

- Divide products into N equal subsets, $P_1, \ldots, P_N$.
- Pick subset P_n and estimate grouping structure and grouping parameters in the other $N-1$ parts.
- Classify products across estimated groups in part P_n and **compute out-of-sample MSE**

$$MSE_n(k) = \frac{1}{J \cdot M} \sum_{m=1}^{M} \sum_{j \in P_n} (y_j - \beta_{m,-n}^{k(j)} x_j - \lambda_{m,-n}^{k(j)})^2$$

- Take average across N folds:

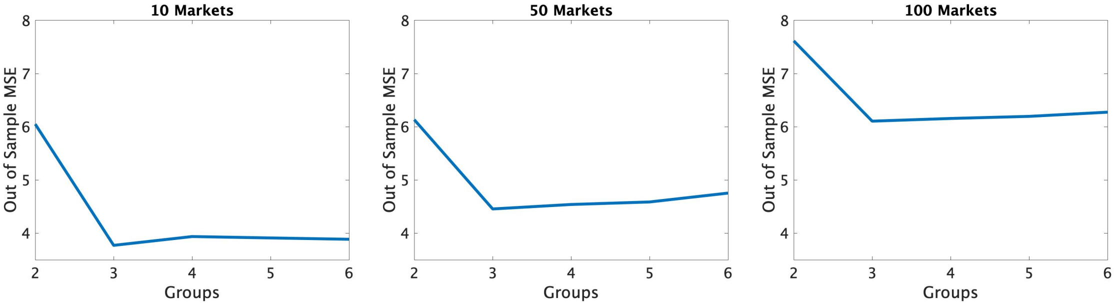
$$MSE(k) = \frac{1}{N} \sum_{n=1}^{N} MSE_n(k)$$

- Choose k according to

$$k^* = \{k(j) | \text{where slope of } MSE(k) \text{ changes}\}$$

Cross validation: Results

K = 3, J = 100, N = 5



Outline

1. Introduction
2. Empirical model
    2.1 Identification
    2.2 Estimation
3. Statistical Properties
4. Monte Carlo
5. Extensions
    5.1 Choosing the number of groups
    5.2 Endogenous prices
    5.3 Type II EV error
    5.4 More flexible empirical models
6. Application: US Automobile Data
7. Conclusion and next steps

Endogenous prices

So far, we have assumed $\mathbb{E}[\xi_{jm} p_{jm}] = 0$.

Endogenous prices

So far, we have assumed $\mathbb{E}[\xi_{jm} p_{jm}] = 0$.

Relax that and assume z_{jm} is a valid instrument for p_{jm} so that $\mathbb{E}[\xi_{jm} | z_{jm}, x_{jm}] = 0$.

Endogenous prices

So far, we have assumed $\mathbb{E}[\xi_{jm} p_{jm}] = 0$.

Relax that and assume z_{jm} is a valid instrument for p_{jm} so that $\mathbb{E}[\xi_{jm} | z_{jm}, x_{jm}] = 0$.

In this case, classify on x_{jm} and z_{jm}:

1. For $(\beta^{1,s}, \dots, \beta^{K,s}, \lambda_1^{K,s}, \dots, \lambda_M^{K,s})$, compute for all $j \in J$:

$$k(j)^{s+1} = \underset{k \in \{1,\dots,K\}}{\arg\min} \sum_{m=1}^{M} \left(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_z^{k,s} z_{jm} + \beta_x^{k,s} x_{jm} + \lambda_m^{k,s}) \right)^2$$

2. Compute:

$$\underset{\beta^1,\dots,\beta^K,\lambda_1^1,\dots,\lambda_M^K}{\arg\min} \sum_{j=1}^{J} \sum_{m=1}^{M} \left(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_z^{k(j),s+1} z_{jm} + \beta_x^{k(j),s+1} x_{jm} + \lambda_m^{k(j),s+1}) \right)^2$$

Endogenous prices

So far, we have assumed $\mathbb{E}[\xi_{jm} p_{jm}] = 0$.

Relax that and assume z_{jm} is a valid instrument for p_{jm} so that $\mathbb{E}[\xi_{jm} | z_{jm}, x_{jm}] = 0$.

In this case, classify on x_{jm} and z_{jm}:

1. For $(\beta^{1,s}, \dots, \beta^{K,s}, \lambda_1^{K,s}, \dots, \lambda_M^{K,s})$, compute for all $j \in J$:

$$k(j)^{s+1} = \underset{k \in \{1,\dots,K\}}{\arg\min} \sum_{m=1}^{M} \left(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_z^{k,s} z_{jm} + \beta_x^{k,s} x_{jm} + \lambda_m^{k,s}) \right)^2$$

2. Compute:

$$\underset{\beta^1,\dots,\beta^K,\lambda_1^1,\dots,\lambda_M^K}{\arg\min} \sum_{j=1}^{J} \sum_{m=1}^{M} \left(\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} - (\beta_z^{k(j),s+1} z_{jm} + \beta_x^{k(j),s+1} x_{jm} + \lambda_m^{k(j),s+1}) \right)^2$$

3. Repeat until convergence of parameters.

Results

Results of 50 Bootstrap iterations

					β_p	β_c	σ_1	σ_2	σ_3
J	M	Runtime	Matched	True	-1	1	0.2	0.3	0.6
100	10	00:02	0.964	Mean β	-0.956	0.956	0.17	0.28	0.596
				Std β	0.052	0.064	0.047	0.056	0.02
100	50	00:15	0.989	Mean β	-0.998	0.999	0.197	0.296	0.596
				Std β	0.015	0.019	0.021	0.029	0.03
100	100	01:45	1.0	Mean β	-1.0	1.001	0.2	0.3	0.6
				Std β	0.007	0.005	0.001	0.001	0.002
500	10	00:16	0.993	Mean β	-0.994	0.995	0.195	0.299	0.6
				Std β	0.016	0.017	0.01	0.008	0.003
500	50	01:38	1.0	Mean β	-1.0	1.001	0.2	0.3	0.6
				Std β	0.004	0.005	0.0	0.001	0.001
500	100	08:01	1.0	Mean β	-1.0	1.0	0.2	0.3	0.6
				Std β	0.003	0.003	0.0	0.0	0.001
1000	10	00:29	0.996	Mean β	-0.993	0.998	0.196	0.3	0.6
				Std β	0.012	0.012	0.006	0.002	0.002
1000	50	05:11	0.986	Mean β	-1.001	1.0	0.198	0.297	0.595
				Std β	0.003	0.003	0.013	0.02	0.033
1000	100	16:15	0.962	Mean β	-0.997	0.997	0.198	0.293	0.591
				Std β	0.015	0.014	0.01	0.037	0.044

Outline

Type II EV error

Assume indirect utility is given by

$$V_{ijm} = \delta_{jm} + \varepsilon_{ijm},$$

where ε_{ijm} is distributed Type II EV with some nesting structure given by $B_1, \ldots, B_K$.

Type II EV error

Assume indirect utility is given by

$$V_{ijm} = \delta_{jm} + \varepsilon_{ijm},$$

where ε_{ijm} is distributed Type II EV with some nesting structure given by $B_1, \ldots, B_K$.

Choice probabilities are given by:

$$\mathbb{P}_{jm} = \frac{\delta_{jm}^{\frac{\theta}{\sigma_k}}}{\sum_j \delta_{jm}^{\frac{\theta}{\sigma_k}}} \frac{\left(\sum_{d \in B_{k(j)}} \delta_{jm}^{\frac{\theta}{\sigma_k}}\right)^{\sigma^k}}{\sum_k \left(\sum_{d \in B_{k(j)}} \delta_{jm}^{\frac{\theta}{\sigma_k}}\right)^{\sigma^k}}.$$

Then, nest can be recovered solving the following problem:

$$\underset{\substack{k(1),\ldots,k(J) \\ \gamma,\beta^1,\ldots,\beta^K,\lambda_1^1,\ldots,\lambda_M^K}}{\arg\min} \sum_{m=1}^{M} \sum_{j=1}^{J} \left(\log \mathbb{P}_{jm} - \beta^k \log(\delta_{jm}(\gamma)) - \lambda_m^k\right)^2$$

Outline

Individual heterogeneity with observed conditional shares

Denote individual heterogeneity by $\omega \sim G(\omega)$

$$\delta_{jm}(\omega) = (\beta_p + \beta_p(\omega))p_{jm} + \beta_x x_{jm} + \xi_{jm}$$

Individual heterogeneity with observed conditional shares

Denote individual heterogeneity by $\omega \sim G(\omega)$

$$\delta_{jm}(\omega) = (\beta_p + \beta_p(\omega))p_{jm} + \beta_x x_{jm} + \xi_{jm}$$

If $G(\omega)$ discrete over types ω^t and $\mathbb{P}^t_{jm} = \mathbb{P}_m(j|\omega^t)$ observed, then:

$$\log \mathbb{P}^t_{jm} - \log \mathbb{P}^t_{0m} = \frac{1}{\sigma_{k^t(j)}}\delta^t_{jm} + (\sigma_{k^t(j)} - 1)\log IV^t_{k^t(j),m}$$

so can classify even type-by-type.

Individual heterogeneity with observed conditional shares

Denote individual heterogeneity by $\omega \sim G(\omega)$

$$\delta_{jm}(\omega) = (\beta_p + \beta_p(\omega))p_{jm} + \beta_x x_{jm} + \xi_{jm}$$

If $G(\omega)$ discrete over types ω^t and $\mathbb{P}^t_{jm} = \mathbb{P}_m(j|\omega^t)$ observed, then:

$$\log \mathbb{P}^t_{jm} - \log \mathbb{P}^t_{0m} = \frac{1}{\sigma_{k^t(j)}}\delta^t_{jm} + (\sigma_{k^t(j)} - 1)\log IV^t_{k^t(j),m}$$

so can classify even type-by-type.

Can also solve joint problem across types to impose constraints, such as $\sigma_{k^t(j)} = \sigma_{k(j)}$ for all t.

Individual unobserved heterogeneity

If $G(\omega)$ discrete over types ω^t and $\mathbb{P}_{jm}(\omega^t) = \mathbb{P}_m(j|\omega^t)$ not observed, then:

$$\underset{\substack{k(1),\ldots,k(J) \\ \beta^1,\ldots,\beta^K,\lambda_1^1,\ldots,\lambda_M^K}}{\arg\min} \sum_{m=1}^{M} \sum_{j=1}^{J} \left(\mathbb{P}_{jm} - \sum_{t=1}^{T} \mathbb{P}(\omega = \omega^t)\mathbb{P}_{jm}(\omega^t)\right)^2, \tag{2}$$

where

$$\mathbb{P}_{jm}(\omega^t) = \frac{e^{\frac{\beta\delta_{jm}(\omega^t)}{\sigma_{k(j)}}} \left(\sum_{d \in B_{k(j)}} e^{\frac{\beta\delta_{dm}(\omega^t)}{\sigma_{k(j)}}}\right)^{\sigma_{k(j)}-1}}{\sum_{k'=1}^{K} \left(\sum_{d \in B_{k'}} e^{\frac{\beta\delta_{dm}(\omega^t)}{\sigma_{k'}}}\right)^{\sigma_{k'}}}.$$

Caveat: requires non-linear optimization!

Next: extend algorithm to BLP contraction?

Higher-order nesting structure

Assume we have upstream nests given by $A_1, \ldots, A_N$, and downstream nests given by $B_1, \ldots, B_K$.

In this case, choice probabilities can be written as:

$$\mathbb{P}_{jm} = \mathbb{P}_{j|k(j),m}\mathbb{P}_{k(j)|n(k),m}\mathbb{P}_{n(k),m},$$

Taking logs, it follows:

$$\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} = \beta^k x_{jm} + (\sigma_k - 1)\ln IV_{km} + (\sigma_n - 1)\log IV_{nm}$$

so that

$$\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}} = \beta^k x_{jm} + \lambda_{k,m}$$

$\implies$ classify using same algorithm to get lower nest structure $\hat{B}_1, \ldots, \hat{B}_K$ and parameters $\hat{\beta}^k, \hat{\lambda}_{k,m}$

Higher-order nesting structure

Given $\hat{B}_1, \ldots, \hat{B}_K$, run modified version of Berry (1994):

$$\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{j'm}} = \beta(x_{jm} - x_{j'm}) + (\sigma^k - 1) \log \frac{\mathbb{P}_{j|k,m}}{\mathbb{P}_{j'|k,m}},$$

where $j, j' \in \hat{B}_k \implies$ recover $\hat{\beta}$ and $\hat{\sigma}_k \implies$ construct plug-in estimation of $(\sigma_k - 1) \log IV_{k,m}$

Higher-order nesting structure

Given $\hat{B}_1, \ldots, \hat{B}_K$, run modified version of Berry (1994):

$$\log \frac{\mathbb{P}_{jm}}{\mathbb{P}_{j'm}} = \beta(x_{jm} - x_{j'm}) + (\sigma^k - 1) \log \frac{\mathbb{P}_{j|k,m}}{\mathbb{P}_{j'|k,m}},$$

where $j, j' \in \hat{B}_k \implies$ recover $\hat{\beta}$ and $\hat{\sigma}_k \implies$ construct plug-in estimation of $(\sigma_k - 1) \log IV_{k,m}$

Recall

$$\lambda_{k,m} = (\sigma_k - 1) \log IV_{k,m} + (\sigma_{n(k)} - 1) \log IV_{n(k),m},$$

$\implies$ can run k-means clustering on $\hat{\lambda}_{k,m} - (\hat{\sigma}_k - 1) \log \hat{IV}_{k,m}$ to recover groups A_n

Application: US Automobile Data

US Automobile data

We use US Automobile data from BLP (1995)
$\longrightarrow$ Data available from `R`-package `hdm` developed by Chernozhukov, Hansen & Spindler (2019)

Information on (essentially) all models marketed between 1971 and 1990

Total sample size is 2217 model/years representing 557 distinct models

We set different years as different markets

Panel construction

Models both enter and exit over this period $\implies$ unbalanced panel

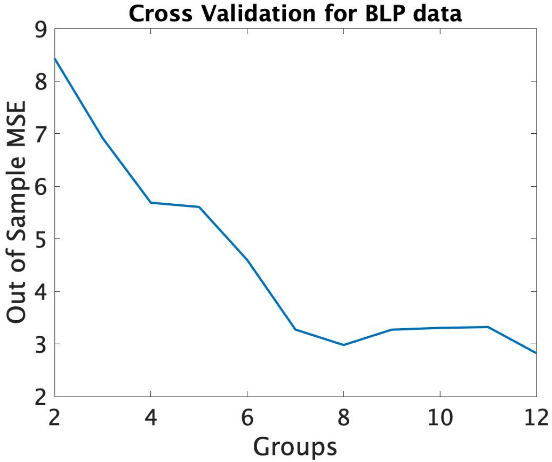
Restrict panel to cars with:

- At least five years of data
- At least three consecutive years
- We are left with 82 products

We adapt our classification algorithm to allow for "missing data":

- Products can **enter** and **exit** over time
- Group of products can also **enter** and **exit** over time!

BLP Application: Choosing the number of groups



BLP Application: First-step Group Characteristics

	Mean	Std.	1	2	3	4	5	6	7	8
Shares	0.001	0.001	0.004	0.009	0.008	0.012	0.006	0.002	0.006	0.002
Price	-0.741	6.898	-3.679	-3.077	-1.694	-1.621	-0.688	-0.610	-0.292	0.211
Log HP	-0.940	0.183	-1.054	-0.973	-0.984	-0.976	-0.942	-0.876	-0.953	-0.915
Log Miles per $	0.767	0.320	0.919	0.623	0.653	0.650	0.823	0.641	0.610	0.642
AC	0.277	0.448	0.072	0.315	0.259	0.268	0.132	0.144	0.303	0.267
Log Space	0.239	0.164	0.096	0.315	0.259	0.282	0.176	0.180	0.303	0.281
Type of car			Subcomp.	Family affordable	Mid-size	Midsize Premium	Compact Luxury	Sport	Family Luxury	Full-size Luxury
# Products	82		7	11	11	15	12	8	12	6

BLP Application: Second-step Results

Estimates Preference Parameters

	$\hat{\beta}$	$\sigma_{\hat{\beta}}$
Price	-0.064***	(0.029)
Horse Power	-0.148	(0.176)
Miles per $	0.222	(0.187)
AC	0.162	(0.133)
Space	0.791	(0.775)

Estimates Within-Nest Correlation

	Group							
	1	2	3	4	5	6	7	8
$\hat{\sigma}$	0.868***	0.596***	0.472***	0.827***	0.722***	0.836***	0.528***	0.572***
$\sigma_{\hat{\sigma}}$	(0.155)	(0.277)	(0.165)	(0.104)	(0.273)	(0.139)	(0.145)	(0.173)
F 1st stage	50.673	2.7697	6.241	6.320	6.963	16.311	11.805	11.748

Conclusion and next steps

Much to do ahead

- Proof of asymptotic consistency for empirical model extensions
- Revisit Monte Carlo with more empirically relevant models
- Empirical applications:
  - IO: Currently working on Nielsen data focusing on ready-to-drink beverages
  - Spatial:
    - Labor markets clusters: Is NYC a closer substitute to SF or Newark?
    - Defining market structure for spatial applications: what's a neighborhood?

Consistency of group estimation

Two key assumptions

– **Group separation**. For simplicity, assume simplest model:

$$\log \mathbb{P}_{jm} = \lambda^{k(j)} + \xi_{jm}, \qquad \text{with } k \in \{1,2\},\ \lambda^2 > \lambda^1,\ \xi_{jm} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$$

It follows

$$\begin{aligned}\mathbb{P}(\hat{k}(j) = 2 | k(j) = 1) &= \mathbb{P}\Big(\sum_{m=1}^{M} (\lambda^1 + \xi_{jm} - \lambda^2)^2 < \sum_{m=1}^{M} (\lambda^1 + \xi_{jm} - \lambda^1)^2\Big) \\ &= \mathbb{P}(\bar{\xi}_{jm} > \lambda^2 - \lambda^1) = 1 - \Phi\Big(\frac{\sqrt{M}}{2}\big(\lambda^2 - \lambda^1\big)\Big) \xrightarrow[M \to \infty]{} 0\end{aligned}$$

Consistency of group estimation

Two key assumptions

- **Group separation**. For simplicity, assume simplest model:

$$\log \mathbb{P}_{jm} = \lambda^{k(j)} + \xi_{jm}, \qquad \text{with } k \in \{1,2\},\ \lambda^2 > \lambda^1,\ \xi_{jm} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$$

  It follows

$$\begin{aligned}\mathbb{P}(\hat{k}(j) = 2 | k(j) = 1) &= \mathbb{P}\Big(\sum_{m=1}^{M} (\lambda^1 + \xi_{jm} - \lambda^2)^2 < \sum_{m=1}^{M} (\lambda^1 + \xi_{jm} - \lambda^1)^2\Big) \\ &= \mathbb{P}(\bar{\xi}_{jm} > \lambda^2 - \lambda^1) = 1 - \Phi\Big(\frac{\sqrt{M}}{2}\big(\lambda^2 - \lambda^1\big)\Big) \xrightarrow[M\to\infty]{} 0\end{aligned}$$

- **Rank condition**: Within-group variation in x for all groups
  $\implies$ separate β^k from λ^k

Back

One-step group estimation

Can combine steps 1 + 2 by solving the following constrained problem:

$$\underset{\substack{k(1),\ldots,k(J)\\ \beta^1,\ldots,\beta^K,\lambda_1^1,\ldots,\lambda_M^K\\ \beta,\sigma^1,\ldots,\sigma^K}}{\arg\min} \sum_{m=1}^{M}\sum_{j=1}^{J}\left(\log\left(\frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}}\right) - (x_{jm}\beta^k + \lambda_m^k)\right)^2,$$

where

$$\beta^k = \frac{\beta}{\sigma^k} \quad \text{and} \quad \lambda_m^k = (\sigma^k - 1)\log\left(\sum_{d\in B_k} e^{\log\left(\frac{\mathbb{P}_{dm}}{\mathbb{P}_{0m}}\right)-\lambda_m^k}\right),$$

but substantial computational/theoretical burden due to non-linear constraints.

Another option is to use Berry inversion directly:

$$\underset{\substack{k(1),\ldots,k(J),\\ \beta,\sigma^1,\ldots,\sigma^K}}{\arg\min} \sum_{m=1}^{M}\sum_{j=1}^{J}\left(\log\left(\frac{\mathbb{P}_{jm}}{\mathbb{P}_{0m}}\right) - (x_{jm}\beta + (1-\sigma^{k(j)})\log\mathbb{P}_{j,m|k(j)}\right)^2,$$

but need to adapt asymptotics to stochastic regressor that varies with group, $\mathbb{P}_{j,m|k(j)}$.

Back

Statistics

Statistics of subsample of cars (N=82)

	Mean	Std. Dev.	Median	Min	Max	t-stat
Price	-.147	7.911	-2.532	-6.601	43.351	-1.06
Miles per Dollar	2.349	.513	2.376	1.352	3.805	2.78
AC	.299	.409	0	0	1	0.49
Miles per Gallon	2.214	.46	2.195	1.38	3.42	1.45
Space	1.266	.187	1.223	.951	1.711	0.13
Horse Power	.407	.069	.386	.308	.727	-0.23
Market Share	.001	.001	.001	0	.004	0.00
Yearly Observations	9.085	4.264	7	5	20	10.42
Year Entry	1980	5.261	1983	1971	1986	-4.62
Year Exit	1989	.88	1990	1988	1990	20.41

Full sample

Statistics for Full Sample

Table: Average characteristics of all cars, (N = 557)

	Mean	Std. Dev.	Median	Min	Max	t-stat
Price	.862	8.983	-2.516	-8.368	43.351	1.06
Miles per Dollar	2.175	.641	2.094	1.055	6.437	-2.78
AC	.275	.424	0	0	1	-0.49
Miles per gallon	2.133	.552	2.07	1	5.3	-1.45
Space	1.263	.216	1.223	.79	1.888	-0.13
Horse Power	.409	.098	.385	.207	.888	0.23
Market Share	.001	.001	0	0	0.006	0.00
Yearly Observations	3.899	3.857	2	1	20	-10.42
Entry Year	1980	6.511	1981	1971	1990	4.62
Exit Year	1984	6.101	1986	1971	1990	-20.41