BLP: The Berry Levinsohn & Pakes (1995) Estimator

Empirical IO: Bergen, 2021

Anders Munk-Nielsen

December 1st, 2021

Dataset

t	j	Name	p_{jt}	$\mathcal{S}_{jt}$	x_{jt}	Instruments
1	1	UK, Ford	100	4%	...	...
1	2	UK, Volvo	110	5%	...	...
2	1	DE, Ford	95	3%	...	...
2	2	DE, Volvo	108	4%	...	...
...	...	...	...	...	...	...

BLP: Simplest Case

BLP Estimation

- **Inversion:** $\hat{\delta}_t := D^{-1}(\mathcal{S}_t)$
- **Linear regression:**

$$\hat{\delta}_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}.$$

BLP: Price Endogeneity

BLP Estimation

- **Inversion:** $\hat{\delta}_t := D^{-1}(\mathcal{S}_t)$
- **Linear IV:** $\hat{\delta}_{jt}$ on (p_{jt}, x_{jt}) with w_{jt} as instrument for p_{jt}.

$$\hat{\delta}_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}.$$

BLP: Flexible Demand

BLP Estimation

Outer: $\min_\theta g_D(\theta)' W g_D(\theta)$ (GMM criterion)

- **Inversion:** $\hat{\delta}_t := D^{-1}(\mathcal{S}_t, \theta)$ (nested iterative algorithm)
- **Linear IV:** $\hat{\delta}_{jt}$ on (p_{jt}, x_{jt}) with w_{jt} as instrument for p_{jt}.

$$\hat{\delta}_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}.$$

  Store residuals, $\hat{\xi}_{jt}$.

- **Criterion:**

$$g_D(\theta) \equiv \frac{1}{\#} \sum_t \sum_j \hat{\xi}_{jt} Z_{jt}^D$$

BLP Estimation

Outer: $\min_{\theta_2} g_D(\theta_2)' W g_D(\theta_2)$ (GMM criterion)

- **Inversion:** $\hat{\boldsymbol{\delta}}_t := D^{-1}(\mathcal{S}_t, \theta_2)$ (nested iterative algorithm)
- **Linear IV:** $\hat{\delta}_{jt}$ on (p_{jt}, x_{jt}) with w_{jt} as instrument for p_{jt}.

$$\hat{\delta}_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}.$$

  Store residuals, $\hat{\xi}_{jt}$, and $\hat{\theta}_1(\theta_2) = (\hat{\alpha}, \hat{\beta})$.

- **Criterion:**

$$g_D(\theta_2) \equiv \frac{1}{\#} \sum_t \sum_j \hat{\xi}_{jt} Z^D_{jt}$$

BLP Estimation: Adding a Supply Side

BLP Estimation

Outer: $\min_\theta g_D(\theta_2)' W g_D(\theta_2)$

- **Inversion:** $\hat{\delta}_t := D^{-1}(\mathcal{S}_t, \theta_2)$ (nested iterative algorithm)
  - **Linear IV, demand:** $\hat{\delta}_{jt}$ on (p_{jt}, x_{jt}) with w_{jt} as instrument for p_{jt}.

    $$\hat{\delta}_{jt} = \alpha p_{jt} + (x_{jt}, v_{jt})\beta + \xi_{jt}.$$

    Store residuals, $\hat{\xi}_{jt}$, and $\hat{\theta}_1(\theta_2) = (\hat{\alpha}, \hat{\beta})$.
  - **Markups:** $\eta_t = \Delta_t(\theta_2)^{-1} s_t$, giving $\hat{c}_{jt} = p_{jt} - \eta_{jt}$
  - **Linear IV, supply:** c_{jt} on (x_{jt}, w_{jt}):

    $$\hat{c}_{jt} = (x_{jt}, w_{jt})\gamma + \omega_{jt},$$

    Store residuals, $\hat{\omega}_{jt}$, and $\hat{\theta}_3(\theta_2) = (\hat{\gamma})$.
  - **Criterion:**

    $$g(\theta) \equiv \begin{pmatrix} \frac{1}{\#} \sum_t \sum_j \hat{\xi}_{jt} Z_{jt}^D \\ \frac{1}{\#} \sum_t \sum_j \hat{\omega}_{jt} Z_{jt}^S \end{pmatrix}$$

Outline

Demand Estimation

- **Independent OLS:**

$$\log s_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}$$

  - Zero cross-price elasticity

- **Independent OLS:**

$$\log s_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}$$

  - Zero cross-price elasticity

- **AIDS/Translog:**

$$\log s_{jt} = \alpha p_{jt} + x_{jt}\beta + \sum_{k \neq j} \gamma_{jk} p_{kt} + \xi_{jt}$$

  - $(J-1)J$ cross-price elasticities: too data hungry

Demand Estimation

- **Independent OLS:**

$$\log s_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}$$

  - Zero cross-price elasticity

- **AIDS/Translog:**

$$\log s_{jt} = \alpha p_{jt} + x_{jt}\beta + \sum_{k \neq j} \gamma_{jk} p_{kt} + \xi_{jt}$$

  - $(J-1)J$ cross-price elasticities: too data hungry

- **Logit:** in market shares

$$\log s_{jt} = \frac{\exp(\alpha p_{jt} + x_{jt}\beta + \xi_{jt})}{\sum_k \exp(\alpha p_{kt} + x_{kt}\beta + \xi_{jt})}$$

Logit Models

- **Logit:** Solves two issues
    - Bring down number of cross-price elasticities to be estimated,
    - Only market-level data available.

Logit Models

- **Logit:** Solves two issues
  - Bring down number of cross-price elasticities to be estimated,
  - Only market-level data available.
- **Intepretation:** Products are bundles of characteristics
  - Different from e.g. AIDS
  - Altering products or adding new products is simple
    - (often a core counterfactual)

Random Utility Models (no nesting)

Random Utility Model

$$U_{ijt} = u_{ijt} + \epsilon_{ijt}, \quad \epsilon_{ijt} \sim \text{IID Extreme Value}$$

Individual i chooses

$$j^* = \arg\max_{j \in J_t} U_{ijt}.$$

Choice probabilities become

$$\Pr(j]i,t) = \frac{\exp(u_{ijt})}{\sum_{k \in J_t} \exp(u_{ikt})}$$

Random Utility Model

$$U_{ijt} = u_{ijt} + \epsilon_{ijt}, \quad \epsilon_{ijt} \sim \text{IID Extreme Value}$$

Individual i chooses

$$j^* = \arg\max_{j \in J_t} U_{ijt}.$$

Choice probabilities become

$$\Pr(j]i, t) = \frac{\exp(u_{ijt})}{\sum_{k \in J_t} \exp(u_{ikt})}$$

- **BLP** sets

$$u_{ijt} = \alpha p_{jt} + x_{jt}\beta_i + \xi_{jt}.$$

  - Notable restrictions: $\alpha_i = \alpha$, and $\xi_{ijt} = \xi_{jt}$
- **Derivatives**: useful later

$$\nabla \Pr(j) = \Pr(j) \left[\nabla u_{ijt} - \sum_{k \in J_t} \Pr(k) \nabla u_{ikt} \right].$$

Homogeneous logit

Model

$$U_{ijt} = \delta_{jt} + \varepsilon_{ijt}.$$

$$\Rightarrow \Pr(j|t) = \frac{\exp(\delta_{jt})}{\sum_{k \in J_t} \exp(\delta_{kt})}$$

- **Identification:** $\delta_{0t} := 0 \; \forall t$.

Homogeneous logit

Model

$$U_{ijt} = \delta_{jt} + \varepsilon_{ijt}.$$

$$\Rightarrow \Pr(j|t) = \frac{\exp(\delta_{jt})}{\sum_{k \in J_t} \exp(\delta_{kt})}$$

- **Identification:** $\delta_{0t} := 0 \ \forall t$.

- **Inversion** derivation

$$\begin{aligned} \log \Pr(j) - \log \Pr(0) &= \delta_{jt} - \Lambda_t - (\delta_{0t} - \Lambda_t) \\ &= \delta_{jt} - \delta_{0t}. \end{aligned}$$

  - where $\Lambda_t \equiv \log \sum_{k \in J_t} \exp(\delta_{kt})$

Model

$$U_{ijt} = \delta_{jt} + \varepsilon_{ijt}.$$

$$\Rightarrow \Pr(j|t) = \frac{\exp(\delta_{jt})}{\sum_{k \in J_t} \exp(\delta_{kt})}$$

- **Identification:** $\delta_{0t} := 0 \ \forall t.$

- **Inversion** derivation

$$\begin{aligned} \log \Pr(j) - \log \Pr(0) &= \delta_{jt} - \Lambda_t - (\delta_{0t} - \Lambda_t) \\ &= \delta_{jt} - \delta_{0t}. \end{aligned}$$

  - where $\Lambda_t \equiv \log \sum_{k \in J_t} \exp(\delta_{kt})$

Inversion

$$D^{-1}(\mathcal{S}_t) = \log \mathcal{S}_{jt} - \log \mathcal{S}_{0t}.$$

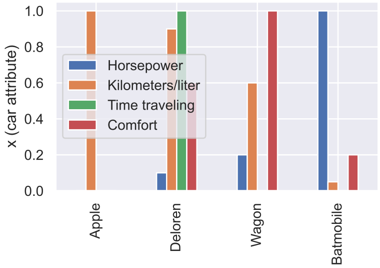
BLP Estimation: Simplest Case

BLP Estimation

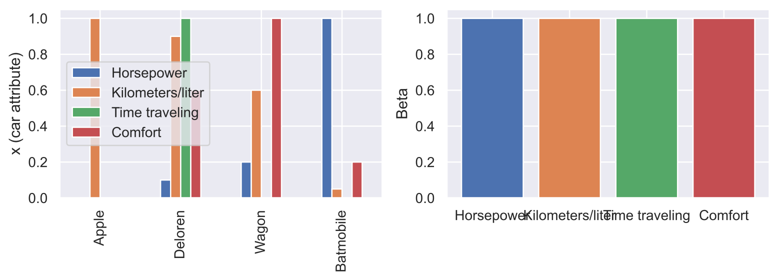
- **Inversion:** Compute $\hat{\delta}_t = D^{-1}(\mathcal{S}_t)$
- **Linear regression** of $\hat{\delta}_t$ on (p_{jt}, x_{jt}):

$$\hat{\delta}_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}.$$

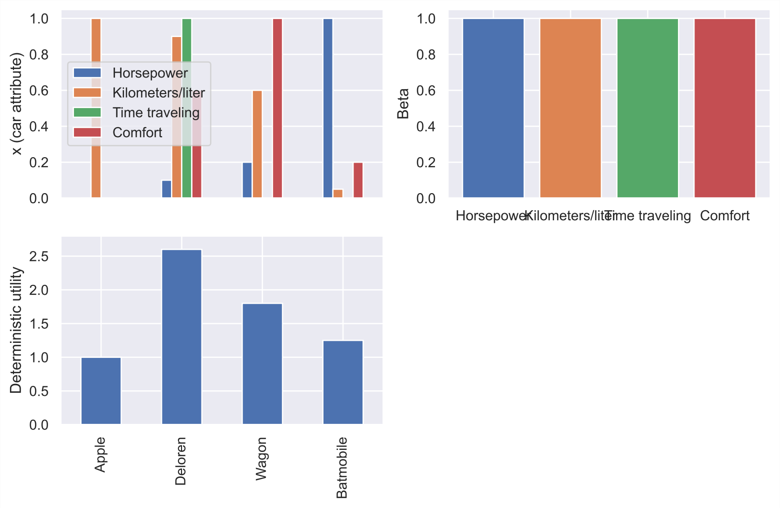
Logit intuition



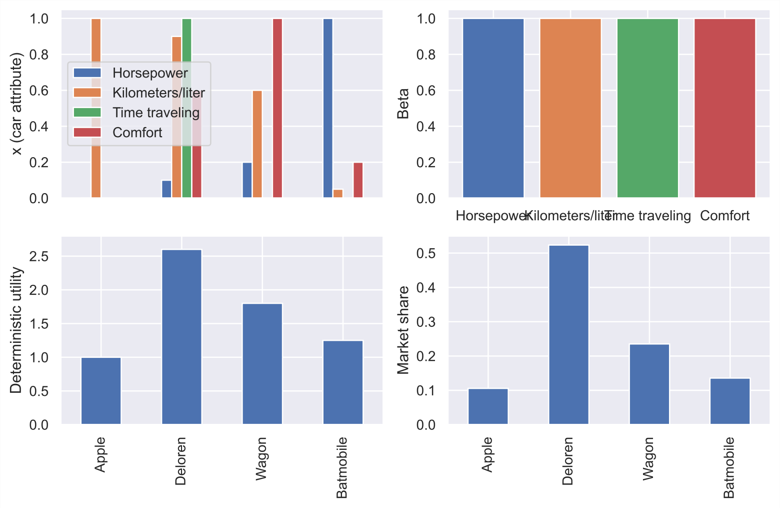
Logit intuition



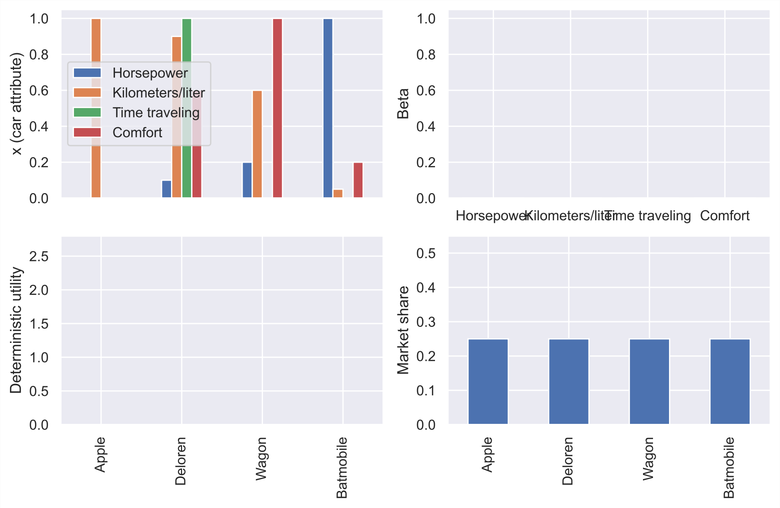
Logit intuition



Logit intuition

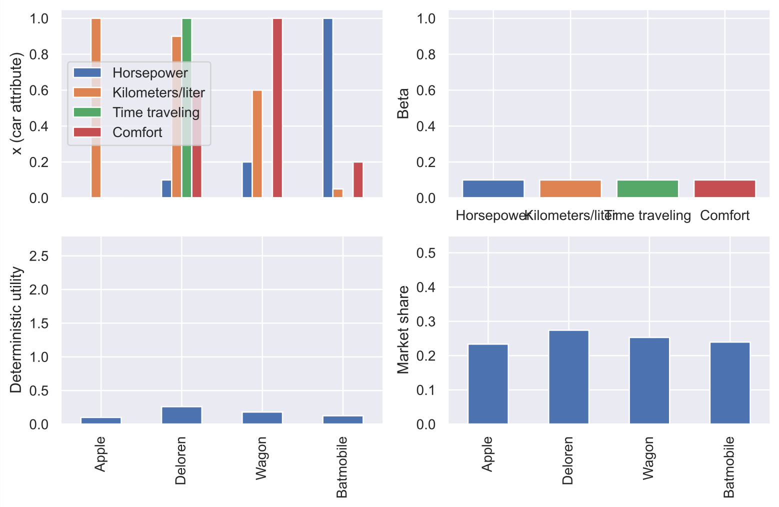


Logit intuition: $\beta = 0$



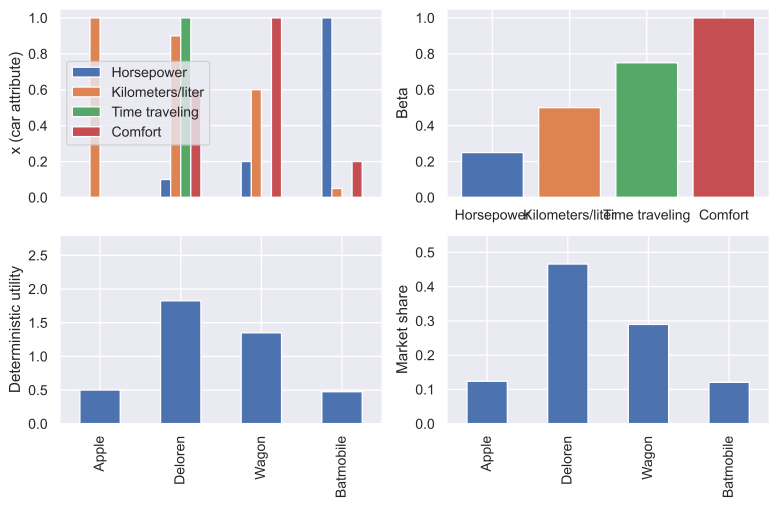
No utility from any characteristics $\Rightarrow$ identical market shares.

Logit intuition: β "small"



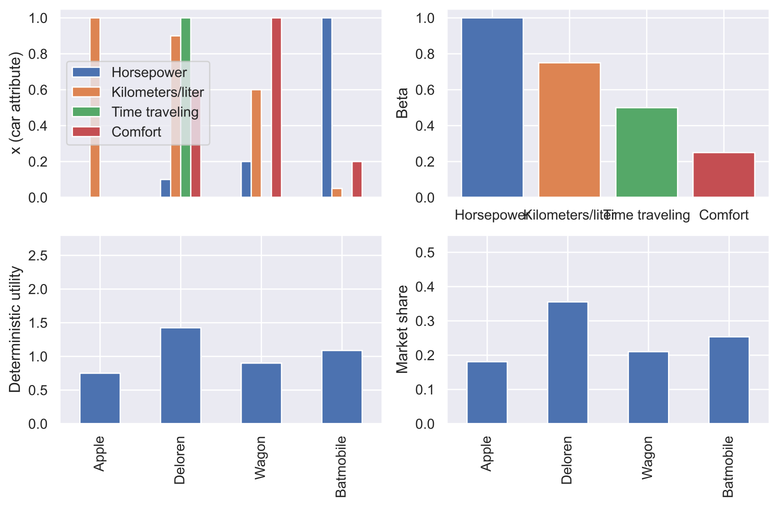
Here, $\beta = (\frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$

Logit intuition



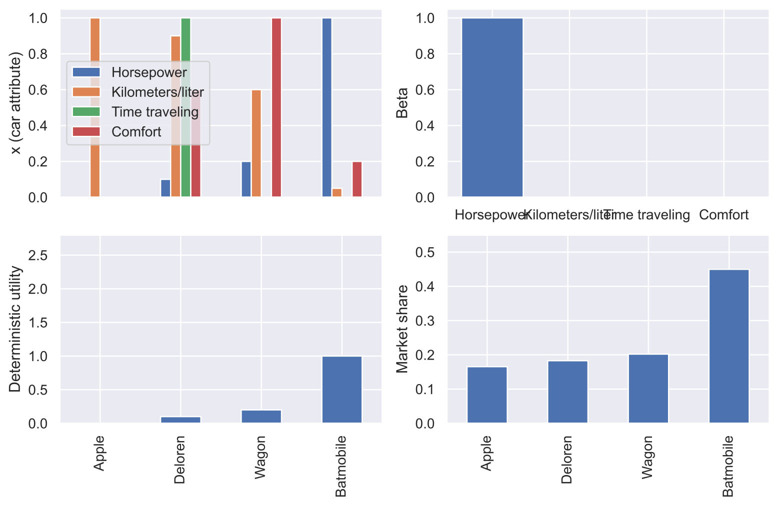
Consumers prefer comfort and time traveling capabilities

Logit intuition



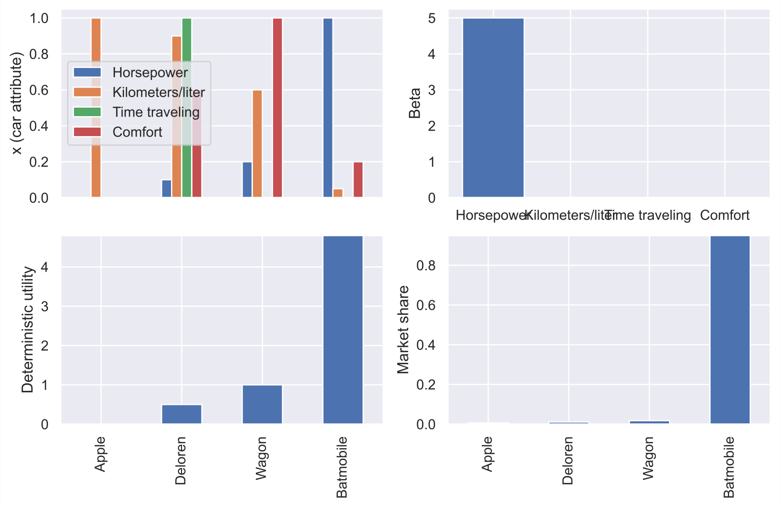
Consumers prefer horsepower and fuel economy.

Logit intuition: $\beta = (1, 0, 0, 0)$



Consumers only care about horsepower (and idiosyncraticies)

Logit intuition: $\beta = (5, 0, 0, 0)$



Consumers only care about horsepower, and by a lot!

Price Endogeneity

- **Problem:** Firms observe ξ_{jt} and set their price accordingly.
  - E.g. ξ_j: Firms observe $\mathbf{1}\{\text{leather interior}\}_j$
  - E.g. ξ_t: Firm prices have seasonality that follows demand
  - E.g. ξ_{jt}: Tesla knows it has become fashionable

BLP Instruments

BLP propose to use $z_{jt} = \sum_{k \neq j} x_{kt}$ as instrument for p_{jt}.

- Captures the "local satiation" of the product space.
- Assumes characteristics are exogenous and prices set subsequently.

Price Endogeneity

- **Problem:** Firms observe ξ_{jt} and set their price accordingly.
  - E.g. ξ_j: Firms observe $\mathbf{1}\{\text{leather interior}\}_j$
  - E.g. ξ_t: Firm prices have seasonality that follows demand
  - E.g. ξ_{jt}: Tesla knows it has become fashionable
- **Instruments:** Price-shifters uncorrelated with demand.
  - Context specific – you have them or you don't.

BLP Instruments

BLP propose to use $z_{jt} = \sum_{k \neq j} x_{kt}$ as instrument for p_{jt}.

- Captures the "local satiation" of the product space.
- Assumes characteristics are exogenous and prices set subsequently.

Price Endogeneity

- **Problem:** Firms observe ξ_{jt} and set their price accordingly.
  - E.g. ξ_j: Firms observe $\mathbf{1}\{\text{leather interior}\}_j$
  - E.g. ξ_t: Firm prices have seasonality that follows demand
  - E.g. ξ_{jt}: Tesla knows it has become fashionable
- **Instruments:** Price-shifters uncorrelated with demand.
  - Context specific – you have them or you don't.
- **Linear IV:** Since the equation is *linear*

$$\delta_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt},$$

  - we can use linear IV.

BLP Instruments

BLP propose to use $z_{jt} = \sum_{k \neq j} x_{kt}$ as instrument for p_{jt}.

- Captures the "local satiation" of the product space.
- Assumes characteristics are exogenous and prices set subsequently.

BLP: Estimation

- **Inversion:** $\hat{\delta}_t := D^{-1}(\mathcal{S}_t)$
- **Linear IV:** $\hat{\delta}_{jt}$ on (p_{jt}, x_{jt}) with w_{jt} as instrument for p_{jt}.

$$\hat{\delta}_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}.$$

Identification

Model

$$
\begin{aligned}
U_{ijt} &= u_{ijt} + \varepsilon_{ijt}, \quad \varepsilon_{ijt} \sim \text{IID Extreme Value T1} \\
\Rightarrow \Pr(j|i,t) &= \frac{\exp(u_{ijt})}{\sum_{k \in J_t} \exp(u_{ikt})}.
\end{aligned}
$$

Model

$$
\begin{aligned}
U_{ijt} &= u_{ijt} + \varepsilon_{ijt}, \quad \varepsilon_{ijt} \sim \text{IID Extreme Value T1} \\
\Rightarrow \Pr(j|i,t) &= \frac{\exp(u_{ijt})}{\sum_{k \in J_t} \exp(u_{ikt})}.
\end{aligned}
$$

- **Result:** CCPs (conditional choice probabilities) are analytically identical under additive rescaling:

$$
\frac{\exp(u_{ijt})}{\sum_{k \in J_t} \exp(u_{ikt})} = \frac{\exp(u_{ijt} - K_{it})}{\sum_{k \in J_t} \exp(u_{ikt} - K_{it})} \quad \forall K_{it} \in \mathbb{R}.
$$

Identification

Model

$$\begin{aligned} U_{ijt} &= u_{ijt} + \varepsilon_{ijt}, \quad \varepsilon_{ijt} \sim \text{IID Extreme Value T1} \\ \Rightarrow \Pr(j|i,t) &= \frac{\exp(u_{ijt})}{\sum_{k \in J_t} \exp(u_{ikt})}. \end{aligned}$$

- **Result:** CCPs (conditional choice probabilities) are analytically identical under additive rescaling:

$$\frac{\exp(u_{ijt})}{\sum_{k \in J_t} \exp(u_{ikt})} = \frac{\exp(u_{ijt} - K_{it})}{\sum_{k \in J_t} \exp(u_{ikt} - K_{it})} \quad \forall K_{it} \in \mathbb{R}.$$

- **Max-rescaling:** Since $\exp(\cdot)$ is numerically unstable, it is *essential* to **avoid overflow** errors:

$$K_{it} := \max_{j \in J_t} u_{ijt}.$$

  - This way, $\exp(u_{ijt} - K_{it}) \leq \exp(0)$, so we will only encounter **underflow**.
  - ... and underflows will cause far fewer issues in estimation

Model

$$
\begin{aligned}
U_{ijt} &= u_{ijt} + \varepsilon_{ijt}, \quad \varepsilon_{ijt} \sim \text{IID Extreme Value T1} \\
\Rightarrow \Pr(j|i,t) &= \frac{\exp(u_{ijt})}{\sum_{k\in J_t} \exp(u_{ikt})}.
\end{aligned}
$$

- **Result:** CCPs (conditional choice probabilities) are analytically identical under additive rescaling:
$$
\frac{\exp(u_{ijt})}{\sum_{k\in J_t} \exp(u_{ikt})} = \frac{\exp(u_{ijt} - K_{it})}{\sum_{k\in J_t} \exp(u_{ikt} - K_{it})} \quad \forall K_{it} \in \mathbb{R}.
$$
- **Max-rescaling:** Since $\exp(\cdot)$ is numerically unstable, it is *essential* to **avoid overflow** errors:
$$
K_{it} := \max_{j\in J_t} u_{ijt}.
$$
    - This way, $\exp(u_{ijt} - K_{it}) \leq \exp(0)$, so we will only encounter **underflow**.
    - ... and underflows will cause far fewer issues in estimation
- **Implication for inversion:** With $u_{ijt} = \delta_{jt}$, we can add any scalar,

Independence of Irrelevant Alternatives (IIA)

- **If** $u_{ijt} = v_{ijt} + \epsilon_{ijt}$, let $\Lambda_{it} \equiv \log \sum_{k \in J_t} \exp(v_{ikt})$

$$\frac{s_{ijt}}{s_{ikt}} = \frac{\exp(v_{ijt})/\Lambda_{it}}{\exp(v_{ikt})/\Lambda_{it}} = \frac{\exp(v_{ijt})}{\exp(v_{ikt})}.$$

  - **Implication:** *relative* CCPs are independent of competing products.
  - If $s_{Tesla} = s_{Leaf}$, then introducing a Porche electric will steal the same market share from the two

Independence of Irrelevant Alternatives (IIA)

- **If** $u_{ijt} = v_{ijt} + \epsilon_{ijt}$, let $\Lambda_{it} \equiv \log \sum_{k \in J_t} \exp(v_{ikt})$

$$\frac{s_{ijt}}{s_{ikt}} = \frac{\exp(v_{ijt})/\Lambda_{it}}{\exp(v_{ikt})/\Lambda_{it}} = \frac{\exp(v_{ijt})}{\exp(v_{ikt})}.$$

    - **Implication:** *relative* CCPs are independent of competing products.
    - If $s_{Tesla} = s_{Leaf}$, then introducing a Porche electric will steal the same market share from the two
- **Cross-price elasticities** will be identical:
    - Similarly: $\frac{\partial s_{Tesla}}{\partial p_{Porsche}} = \frac{\partial s_{Tesla}}{\partial p_{Leaf}}$ (if $s_{Porsche} = s_{Leaf}$)

Independence of Irrelevant Alternatives (IIA)

- **If** $u_{ijt} = v_{ijt} + \epsilon_{ijt}$, let $\Lambda_{it} \equiv \log \sum_{k \in J_t} \exp(v_{ikt})$

$$\frac{s_{ijt}}{s_{ikt}} = \frac{\exp(v_{ijt})/\Lambda_{it}}{\exp(v_{ikt})/\Lambda_{it}} = \frac{\exp(v_{ijt})}{\exp(v_{ikt})}.$$

  - **Implication:** *relative* CCPs are independent of competing products.
  - If $s_{Tesla} = s_{Leaf}$, then introducing a Porche electric will steal the same market share from the two
- **Cross-price elasticities** will be identical:
  - Similarly: $\frac{\partial s_{Tesla}}{\partial p_{Porsche}} = \frac{\partial s_{Tesla}}{\partial p_{Leaf}}$ (if $s_{Porsche} = s_{Leaf}$)
- **Beating IIA:** two common approaches

Independence of Irrelevant Alternatives (IIA)

- **If** $u_{ijt} = v_{ijt} + \epsilon_{ijt}$, let $\Lambda_{it} \equiv \log \sum_{k \in J_t} \exp(v_{ikt})$

$$\frac{s_{ijt}}{s_{ikt}} = \frac{\exp(v_{ijt})/\Lambda_{it}}{\exp(v_{ikt})/\Lambda_{it}} = \frac{\exp(v_{ijt})}{\exp(v_{ikt})}.$$

  - **Implication:** *relative* CCPs are independent of competing products.
  - If $s_{Tesla} = s_{Leaf}$, then introducing a Porche electric will steal the same market share from the two
- **Cross-price elasticities** will be identical:
  - Similarly: $\frac{\partial s_{Tesla}}{\partial p_{Porsche}} = \frac{\partial s_{Tesla}}{\partial p_{Leaf}}$ (if $s_{Porsche} = s_{Leaf}$)
- **Beating IIA:** two common approaches
- **Random coefficients:** the *aggregate* market share function will not suffer from IIA.
  - If $s_{iLeaf} \cong 0, s_{iTesla} \gg 0$ for rich i, then Tesla grows more, Vise versa for poor i.

Solving IIA

- **Two approaches**
  1. Random coefficients logit (RCL)
  2. Nested logit (NL)
  3. (both, RCNL)

Solving IIA

- **Two approaches**
  1. Random coefficients logit (RCL)
  2. Nested logit (NL)
  3. (both, RCNL)
- **1 RCL:** Consumers are heterogeneous
  - But IIA persists *within* a given consumer...
  - $\Rightarrow$ ... i.e. consumer i has flat cross-price elasticities

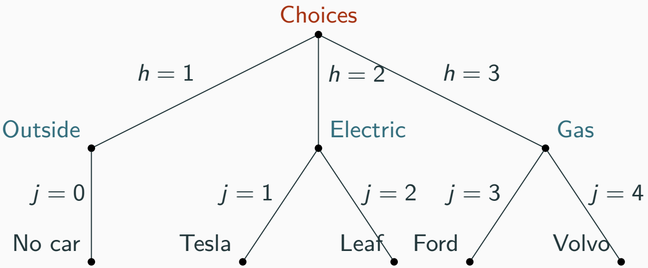
Solving IIA

- **Two approaches**
  1. Random coefficients logit (RCL)
  2. Nested logit (NL)
  3. (both, RCNL)
- **1 RCL:** Consumers are heterogeneous
  - But IIA persists *within* a given consumer...
  - $\Rightarrow$ ... i.e. consumer i has flat cross-price elasticities
- **2 NL:** Error terms are correlated within a nest
  - Still IIA within a *nest*
    - But different price elasticities across nests
  - But elasticities are identical across consumers

Solving IIA

- **Two approaches**
  1. Random coefficients logit (RCL)
  2. Nested logit (NL)
  3. (both, RCNL)
- **1 RCL:** Consumers are heterogeneous
  - But IIA persists *within* a given consumer...
  - $\Rightarrow$ ... i.e. consumer i has flat cross-price elasticities
- **2 NL:** Error terms are correlated within a nest
  - Still IIA within a *nest*
    - But different price elasticities across nests
  - But elasticities are identical across consumers
- **Micro vs. macro data:** not clear precisely what empirically identifies the two approaches separately with only *aggregate* share data

Nested Logit



Example with 3 nests:

1. Outside option
2. Electric cars
3. Gasoline cars

Nested Logit

$$U_{ijt} = u_{ijt} + \epsilon_{ijt},$$

where products are organized into a partition of *nests*, $h_j \in H = \{1, ..., H\}$. Then Choice probabilities become

$$\begin{aligned} \Pr(j) &= \Pr(\text{nest } h_j) \Pr(j|\text{nest } h_j), \\ \text{where} \quad \Pr(\text{nest } h_j) &= \frac{\exp(I_{ih_jt})}{\sum_{h \in H} \exp(I_{iht})} \\ \Pr(j|\text{nest } h) &= \frac{\exp(\frac{1}{1-\rho} u_{ijt})}{\exp(\frac{1}{1-\rho} I_{iht})}, \end{aligned}$$

and the *inclusive value* (or "logsum") is

$$I_{iht} \equiv (1-\rho) \log \sum_{j \in J_{ht}} \exp\left(\frac{1}{1-\rho} u_{ijt}\right)$$

Inversion with Nested Logit

- **Turns out** inversion can be solved *analytically*
  - Holds only for the simplest vanilla nested logit
- **Model**

$$
\begin{aligned}
U_{ijt} &= \delta_{jt} + \epsilon_{ijt} \\
\Pr(j|t) &= \frac{\exp(IV_h)}{\sum_{h' \in H} \exp(IV_{h'})} \frac{\exp(\tilde{\delta}_{jt})}{\sum_{k \in J_{ht}} \exp(\tilde{\delta}_{kt})}
\end{aligned}
$$

Analytic Inversion (Berry, 1994)

$$D^{-1}(\mathcal{S}_t) = \log \mathcal{S}_{jt} - \log \mathcal{S}_{0t} - \rho \log \mathcal{S}_{j|ht}$$

where $\mathcal{S}_{j|ht}$ is the market share of j in nest h.

BLP: Random Coefficients

Outer: $\min_\theta g_D(\theta)' W g_D(\theta)$ (GMM criterion)

- **Inversion:** $\hat{\delta}_t := D^{-1}(\mathcal{S}_t, \theta)$ (nested iterative algorithm)
- **Linear IV:** $\hat{\delta}_{jt}$ on (p_{jt}, x_{jt}) with w_{jt} as instrument for p_{jt}.

$$\hat{\delta}_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}.$$

  Store residuals, $\hat{\xi}_{jt}$.
- **Criterion:**

$$g_D(\theta) \equiv \frac{1}{\#} \sum_t \sum_j \hat{\xi}_{jt} Z_{jt}^D$$

BLP Estimation

Outer: $\min_\theta g_D(\theta_2)' W g_D(\theta_2)$ (GMM criterion)

- **Inversion:** $\hat{\delta}_t := D^{-1}(\mathcal{S}_t, \theta_2)$ (nested iterative algorithm)
- **Linear IV:** $\hat{\delta}_{jt}$ on (p_{jt}, x_{jt}) with w_{jt} as instrument for p_{jt}.

$$\hat{\delta}_{jt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt}.$$

  Store residuals, $\hat{\xi}_{jt}$, and $\hat{\theta}_1(\theta_2) = (\hat{\alpha}, \hat{\beta})$.
- **Criterion:**

$$g_D(\theta_2) \equiv \frac{1}{\#} \sum_t \sum_j \hat{\xi}_{jt} Z_{jt}^D$$

Random Coefficients

Random Coefficients Logit

$$\begin{aligned} U_{ijt} &= \alpha p_{jt} + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt}, \\ \Rightarrow \Pr(j|\beta_i, t) &= \frac{\exp(\alpha p_{jt} + x_{jt}\beta_i + \xi_{jt})}{\sum_{k \in J_t} \exp(\alpha p_{kt} + x_{kt}\beta_i + \xi_{kt})}. \end{aligned}$$

Random Coefficients Logit

$$\begin{aligned} U_{ijt} &= \alpha p_{jt} + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt}, \\ \Rightarrow \Pr(j|\beta_i, t) &= \frac{\exp(\alpha p_{jt} + x_{jt}\beta_i + \xi_{jt})}{\sum_{k\in J_t} \exp(\alpha p_{kt} + x_{kt}\beta_i + \xi_{kt})}. \end{aligned}$$

- **IIA** still holds *within individual*
- Recall: $\nabla \Pr(j) = \Pr(j)\left[\nabla u_{ijt} - \sum_{k\in J_t} \Pr(k)\nabla u_{ikt}\right]$,
- ... so since $\frac{\partial u_{ikt}}{\partial p_{i\ell t}} = \alpha \mathbf{1}_{\ell=k}$, the (semi-)elasticity is

$$\frac{\partial \log \Pr_i(j)}{\partial p_\ell} = \alpha \mathbf{1}_{\ell=j} - \Pr_i(\ell)\alpha.$$

  - So *cross*-price elasticities are independent of j!
  - A Tesla steals the same from all cars in i's choiceset
  - ... but the effect differs over i: proportionally to $\Pr_i(j)$.

From individual to market shares

Random Coefficients Logit

$$
\begin{aligned}
U_{ijt} &= \alpha p_{jt} + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt}, \\
\Rightarrow \Pr(j|\beta_i, t) &= \frac{\exp(\alpha p_{jt} + x_{jt}\beta_i + \xi_{jt})}{\sum_{k\in J_t} \exp(\alpha p_{kt} + x_{kt}\beta_i + \xi_{kt})}.
\end{aligned}
$$

From individual to market shares

Random Coefficients Logit

$$\begin{aligned} U_{ijt} &= \alpha p_{jt} + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt}, \\ \Rightarrow \Pr(j|\beta_i, t) &= \frac{\exp(\alpha p_{jt} + x_{jt}\beta_i + \xi_{jt})}{\sum_{k\in J_t} \exp(\alpha p_{kt} + x_{kt}\beta_i + \xi_{kt})}. \end{aligned}$$

- $\Pr_i(j)$ is *individual* i's demand
- We observe (and firms optimize wrt.) the *market share*

$$s_{jt} = \int \Pr(j|\beta, t)\, \mathrm{d}F(\beta) = N^{-1} \sum_{i=1}^{N} \Pr(j|\beta_i, t).$$

- $F(\beta)$ obviously needs to be restricted parametrically. Commonly:

$$\beta_i \sim \mathcal{N}(\beta, \Sigma).$$

Here, $\tilde{\theta}_2$ will parameterize Σ, while $\beta \in \theta_1$.

Mixed Logit

- **Common/idiosyncratic dichotomy:** $u_{ijt} = \delta_{jt} + \mu_{ijt}$

$$s_{jt}(\boldsymbol{\delta}_t, \tilde{\theta}_2) = \int \frac{\exp(\delta_{jt} + \mu_{ijt})}{\sum_{k \in J} \exp(\delta_{kt} + \mu_{ikt})} \, \mathrm{d}F(\boldsymbol{\mu}_{it} | \tilde{\theta}_2)$$

- **Example:** Suppose $\dim(x_{jt}) = 1$, then $\beta_i \sim \mathcal{N}(\beta, \sigma_\beta^2)$

$$\begin{aligned} u_{ijt} &= \alpha p_{jt} + \beta_i x_{jt} + \xi_{jt} \\ &= \underbrace{\alpha p_{jt} + \beta x_{jt} + \xi_{jt}}_{=\delta_{jt}} + \underbrace{\sigma_\beta \nu_i x_{jt}}_{=\mu_{ijt}}, \quad \nu_i \sim \mathcal{N}(0,1) \end{aligned}$$

- **Resulting** market share function

$$s_{jt}(\boldsymbol{\delta}_t, \sigma_\beta) = \int_{-\infty}^{\infty} \frac{\exp(\delta_{jt} + \sigma_\beta \nu x_{jt})}{\sum_{k \in J} \exp(\delta_{kt} + \sigma_\beta \nu x_{kt})} \phi(\nu) \, \mathrm{d}\nu$$

Computing the Integral

Integration Problem

$$\int_{-\infty}^{\infty} \varphi(\nu)\phi(\nu)\,\mathrm{d}\nu$$

where $\varphi(\nu) = \frac{\exp(\delta_{jt}+\sigma_\beta \nu x_{jt})}{\sum_{k\in J}\exp(\delta_{kt}+\sigma_\beta \nu x_{kt})}$.

- **Quadrature**: Given *weights* and *nodes*, $\{w_q, x_q\}_{q=1}^{Q}$

$$\int_{-\infty}^{\infty} \varphi(\nu)\phi(\nu)\,\mathrm{d}\nu \cong \sum_{q=1}^{Q} w_q \varphi(x_q)$$

  - Exactly integrates a Q'th order polynomial approximation of $\varphi(\cdot)$.
  - In higher dimensions: cartesian grids $\Rightarrow$ curse of dimensionality.

Integration Problem

$$\int_{-\infty}^{\infty} \varphi(\nu)\phi(\nu)\,\mathrm{d}\nu$$

where $\varphi(\nu) = \frac{\exp(\delta_{jt}+\sigma_\beta \nu x_{jt})}{\sum_{k\in J}\exp(\delta_{kt}+\sigma_\beta \nu x_{kt})}$.

- **Quadrature**: Given *weights* and *nodes*, $\{w_q, x_q\}_{q=1}^{Q}$

$$\int_{-\infty}^{\infty} \varphi(\nu)\phi(\nu)\,\mathrm{d}\nu \cong \sum_{q=1}^{Q} w_q \varphi(x_q)$$

  - Exactly integrates a Q'th order polynomial approximation of $\varphi(\cdot)$.
  - In higher dimensions: cartesian grids $\Rightarrow$ curse of dimensionality.
- **Simulation:** Draw $\{\nu_r\}_{r=1}^{R}$ with $\nu_r \sim \mathrm{IID}\mathcal{N}(0,1)$:

$$\int_{-\infty}^{\infty} \varphi(\nu)\phi(\nu)\,\mathrm{d}\nu \cong R^{-1}\sum_{r=1}^{R} \varphi(\nu_r)$$

  - Generalizes straightforwardly to $\dim(\beta_i) > 1$

Finding δ: The Nested Fixed Point

- **Challenge:** How to find δ?
- **Berry 1994** (and BLP 1995) showed that $\Gamma : \mathbb{R}^{J_t} \rightarrow \mathbb{R}^{J_t}$ is a contraction:

$$\Gamma(\boldsymbol{\delta}_t) \equiv \boldsymbol{\delta}_t + \log \mathcal{S}_t - \log \mathbf{s}_j(\boldsymbol{\delta}_t, \tilde{\theta}_2).$$

Finding δ: The Nested Fixed Point

- **Challenge:** How to find δ?
- **Berry 1994** (and BLP 1995) showed that $\Gamma : \mathbb{R}^{J_t} \to \mathbb{R}^{J_t}$ is a contraction:

$$\Gamma(\boldsymbol{\delta}_t) \equiv \boldsymbol{\delta}_t + \log \mathcal{S}_t - \log \mathbf{s}_j(\boldsymbol{\delta}_t, \tilde{\theta}_2).$$

FXP by Iterating on Γ

- Initialize $\boldsymbol{\delta}_t^{[0]} := \log \mathcal{S}_t$
- Update $\boldsymbol{\delta}_t^{[i]} := \Gamma(\boldsymbol{\delta}_t^{[i-1]})$
- Stop if $\|\boldsymbol{\delta}_t^{[i]} - \boldsymbol{\delta}_t^{[i-1]}\| < \epsilon^{\text{tol}}$

Adviseable to set $\epsilon^{\text{tol}} < 10^{-14}$.

Finding δ: The Nested Fixed Point

- **Challenge:** How to find δ?
- **Berry 1994** (and BLP 1995) showed that $\Gamma : \mathbb{R}^{J_t} \rightarrow \mathbb{R}^{J_t}$ is a contraction:
$$\Gamma(\boldsymbol{\delta}_t) \equiv \boldsymbol{\delta}_t + \log \mathcal{S}_t - \log \mathbf{s}_j(\boldsymbol{\delta}_t, \tilde{\theta}_2).$$

FXP by Iterating on Γ

- Initialize $\boldsymbol{\delta}_t^{[0]} := \log \mathcal{S}_t$
- Update $\boldsymbol{\delta}_t^{[i]} := \Gamma(\boldsymbol{\delta}_t^{[i-1]})$
- Stop if $\|\boldsymbol{\delta}_t^{[i]} - \boldsymbol{\delta}_t^{[i-1]}\| < \epsilon^{\text{tol}}$

Adviseable to set $\epsilon^{\text{tol}} < 10^{-14}$.

- **Problem:** Linear convergence $\Rightarrow$ expensive to get the final few steps.

Solving for δ

The inversion problem

For each $t \in \{1, ..., T\}$, solve the J_t equations in J_t unknown δ_{jt}s:

$$\mathcal{S}_{jt} = s_{jt}(\boldsymbol{\delta}_t, \theta_2) \, \forall j.$$

Solving for δ

The inversion problem

For each $t \in \{1, ..., T\}$, solve the J_t equations in J_t unknown δ_{jt}s:

$$\mathcal{S}_{jt} = s_{jt}(\boldsymbol{\delta}_t, \theta_2)\,\forall j.$$

Algorithms

- **Direct** iterations on $\Gamma(\boldsymbol{\delta}_t^{[i]}) = \boldsymbol{\delta}_t^{[i]} + \log \mathcal{S}_t - \log \mathbf{s}_t(\boldsymbol{\delta}_t, \tilde{\theta}_2)$
- **Newton-type** method:

$$\boldsymbol{\delta}_t^{[i]} := \boldsymbol{\delta}_t^{[i-1]} - \lambda \left[\nabla \mathbf{s}_t\right]^{-1} \mathbf{s}_t$$

- **SQUAREM**: Avoids the Jacobian, utilizing only $\boldsymbol{\delta}_t^{[i]}, \Gamma(\boldsymbol{\delta}_t^{[i]}), \Gamma(\Gamma(\boldsymbol{\delta}_t^{[i]}))$.

Fixed effects?

- **Common** to have many markets and/or time periods
  - $\Rightarrow$ obvious to want $\xi_{jt} = \xi_j + \xi_t + \xi_{jt}$
- **Challenge:** T and/or J may be so large that LSDV becomes burdensome.

Fixed effects?

- **Common** to have many markets and/or time periods
  - $\Rightarrow$ obvious to want $\xi_{jt} = \xi_j + \xi_t + \xi_{jt}$
- **Challenge:** T and/or J may be so large that LSDV becomes burdensome.

Method of Alternating Projections (MAP)

Initialize Y_{jt} and X_{jt} to data counterparts. Then do

- $Y_{jt} := Y_{jt} - \overline{Y}_j - \overline{Y}_t$
- $X_{jt} := X_{jt} - \overline{X}_j - \overline{X}_t$

Stop if $\overline{X}_j$ and $\overline{X}_t$ are zero.

(Correia, 2016: `reghdfe` and `ivreghdfe` in Stata)

Estimating ξ_{jt}

- **Temptation** if one has micro data: simply estimate JT dummies, ξ_{jt}.

Estimating ξ_{jt}

- **Temptation** if one has micro data: simply estimate JT dummies, ξ_{jt}.
- **Problem:** p_{jt} and x_{jt} have zero variation
  - Centralized market assumption.
  - With individualized prices, there may be a scope...
    - ... but then the assumption of homogenous $\alpha_i = \alpha$ may be unpalatable

Estimating ξ_{jt}

- **Temptation** if one has micro data: simply estimate JT dummies, ξ_{jt}.
- **Problem:** p_{jt} and x_{jt} have zero variation
  - Centralized market assumption.
  - With individualized prices, there may be a scope...
    - ... but then the assumption of homogenous $\alpha_i = \alpha$ may be unpalatable
- **Random effects approach:** May be feasible to specify $F_\xi(\xi|x, w)$ and approach with ML.

Outline

Profit Maximization Problem

The Firm's Problem (single-product case)

$$\max_{p_{jt}} (p_{jt} - c_{jt}) s_{jt}$$

The Firm's Problem (single-product case)

$$\max_{p_{jt}} (p_{jt} - c_{jt}) s_{jt}$$

- **FOC:**

$$\begin{aligned} \Rightarrow 0 &= s_{jt} + (p_{jt} - c_{jt}) \frac{\partial s_{jt}}{\partial p_{jt}} \\ \Leftrightarrow p_{jt} &= c_{jt} + \underbrace{\left(\frac{\partial s_{jt}}{\partial p_{jt}} \right)^{-1} s_{jt}}_{=\eta_{jt}(p_{jt}, s_{jt}, \theta_2)}. \end{aligned}$$

Profit Maximization Problem

The Firm's Problem (single-product case)

$$\max_{p_{jt}} (p_{jt} - c_{jt}) s_{jt}$$

- **FOC:**

$$\begin{aligned} \Rightarrow 0 &= s_{jt} + (p_{jt} - c_{jt}) \frac{\partial s_{jt}}{\partial p_{jt}} \\ \Leftrightarrow p_{jt} &= c_{jt} + \underbrace{\left(\frac{\partial s_{jt}}{\partial p_{jt}} \right)^{-1} s_{jt}}_{= \eta_{jt}(p_{jt}, s_{jt}, \theta_2)}. \end{aligned}$$

- **Parameterizing marginal cost:**

$$p_{jt} - \eta_{jt} = x_{jt}\gamma_1 + w_{jt}\gamma_2 + \omega_{jt}.$$

  - Marginal cost shifters, w_{jt}, are excluded from demand.

Profit Maximization Problem

The Firm's Problem (single-product case)

$$\max_{p_{jt}} (p_{jt} - c_{jt}) s_{jt}$$

- **FOC:**

$$\begin{aligned} \Rightarrow 0 &= s_{jt} + (p_{jt} - c_{jt}) \frac{\partial s_{jt}}{\partial p_{jt}} \\ \Leftrightarrow p_{jt} &= c_{jt} + \underbrace{\left(\frac{\partial s_{jt}}{\partial p_{jt}} \right)^{-1} s_{jt}}_{=\eta_{jt}(p_{jt}, s_{jt}, \theta_2)}. \end{aligned}$$

- **Parameterizing marginal cost:**

$$p_{jt} - \eta_{jt} = x_{jt}\gamma_1 + w_{jt}\gamma_2 + \omega_{jt}.$$

  - Marginal cost shifters, w_{jt}, are excluded from demand.
- **Supply moments:** Assume $\mathbb{E}(\omega_{jt} Z^S_{jt}) = 0$.
  - Criterion contribution: $g^S(\theta) = N^{-1} \sum_{j,t} \hat{\omega}_{jt} Z^S_{jt}$.

Multi-product Firms

The Firm's Problem (multi-product case)

$$\max_{\{p_{jt}:\forall j \in J_{ft}\}} \sum_{j \in J_{ft}} (p_{jt} - c_{jt}) s_{jt}$$

The Firm's Problem (multi-product case)

$$\max_{\{p_{jt}:\forall j\in J_{ft}\}} \sum_{j\in J_{ft}} (p_{jt}-c_{jt})s_{jt}$$

- **FOC** now internalizes cannibalization

$$\text{FOC}: s_{jt}(\mathbf{p}_t) + \sum_{k\in J_{ft}} \frac{\partial s_{kt}}{\partial p_{jt}}(p_{kt}-c_{kt}) = 0$$

  - FOC appropriate since $s_{jt}(p_j, p_{-j})$ is everywhere smooth

- **Stacking** and defining $\Delta_t(\mathbf{p}_t) \equiv -\mathcal{H}_t \odot \nabla_{\mathbf{p}_t}\mathbf{s}_t$ $(J_t \times J_t)$

$$\begin{aligned} \mathbf{s}_t(\mathbf{p}_t) &= \Delta_t(\mathbf{p}_t)(\mathbf{p}_t - \mathbf{c}_t) \\ \Leftrightarrow \mathbf{p}_t &= \mathbf{c}_t + \underbrace{[\Delta_t(\mathbf{p}_t)]^{-1}\mathbf{s}_t(\mathbf{p}_t)}_{=\boldsymbol{\eta}_t(\mathbf{p}_t, s_t, \theta_2)}. \end{aligned}$$

  - $\{\mathcal{H}_t\}_{k,\ell} = \mathbf{1}\{k \text{ and } \ell \text{ produced by same firm}\}$.

Solving for the NE

Problem: Computing the Price Equilibrium

$$\mathbf{p}_t = \mathbf{c}_t + \underbrace{\left[-\mathcal{H}_t \odot \nabla_{\mathbf{p}_t} \mathbf{s}_t\right]^{-1} \mathbf{s}_t(\mathbf{p}_t)}_{=\boldsymbol{\eta}_t(\mathbf{p}_t, s_t, \theta_2)}.$$

Solving for the NE

Problem: Computing the Price Equilibrium

$$\mathbf{p}_t = \mathbf{c}_t + \underbrace{\left[-\mathcal{H}_t \odot \nabla_{\mathbf{p}_t}\mathbf{s}_t\right]^{-1}\mathbf{s}_t(\mathbf{p}_t)}_{=\boldsymbol{\eta}_t(\mathbf{p}_t, s_t, \theta_2)}.$$

- **Iterated Best Response:**

$$\mathbf{p}_t^{[i]} := \mathbf{p}_t^{[i-1]} + \boldsymbol{\eta}_t(\mathbf{p}_t, \mathbf{s}_t, \theta_2).$$

  - NB! Not a guaranteed contraction mapping!
- **Morrow & Skerlow (2011):** Split

$$\nabla_{\mathbf{p}}\mathbf{s}_t = \Upsilon_t^{\text{own}}(\mathbf{p}_t) - \Upsilon_t^{\text{cross}}(\mathbf{p}_t)$$

$$\begin{aligned}
\Upsilon_t^{\text{own}}(\mathbf{p}_t) &= \int \alpha s_{ijt}(\boldsymbol{\mu}_{it}) f(\boldsymbol{\mu}_{it}|\tilde{\theta}_2)\mathrm{d}\boldsymbol{\mu}_{it} \\
\Upsilon_t^{\text{cross}}(\mathbf{p}_t) &= \int \alpha s_{ijt}(\boldsymbol{\mu}_{it}) s_{ikt}(\boldsymbol{\mu}_{it}) f(\boldsymbol{\mu}_{it}|\tilde{\theta}_2)\mathrm{d}\boldsymbol{\mu}_{it}
\end{aligned}$$

and use the fixed point

$$\mathbf{p}_t^{[i]} = \mathbf{c}_t + \Upsilon^{\text{own}}(\mathbf{p}_t^{[i-1]})^{-1}\left[\mathcal{H}_t \odot \Upsilon^{\text{cross}}(\mathbf{p}_t^{[i-1]})\right](\mathbf{p}_t^{[i-1]} - \mathbf{c}_t) - \Upsilon^{\text{own}}\mathbf{s}_t(\mathbf{p}_t^{[i-1]})$$

BLP Estimation

Outer: $\min_\theta g_D(\theta_2)' W g_D(\theta_2)$

- **Inversion:** $\hat{\delta}_t := D^{-1}(\mathcal{S}_t, \theta_2)$ (nested iterative algorithm)
  - **Linear IV, demand:** $\hat{\delta}_{jt}$ on (p_{jt}, x_{jt}) with w_{jt} as instrument for p_{jt}.

    $$\hat{\delta}_{jt} = \alpha p_{jt} + (x_{jt}, v_{jt})\beta + \xi_{jt}.$$

    Store residuals, $\hat{\xi}_{jt}$, and $\hat{\theta}_1(\theta_2) = (\hat{\alpha}, \hat{\beta})$.
  - **Prices:** Solve price equilibrium:

    $$\mathbf{p}_t = \mathbf{c}_t + [-\mathcal{H}_t \odot \nabla \mathbf{s}_t(\mathbf{p}_t)]^{-1} \mathbf{s}_t(\mathbf{p}_t)$$

    yielding $\hat{c}_{jt}$.
  - **Linear IV, supply:** c_{jt} on (x_{jt}, w_{jt}) with v_{jt} as instruments:

    $$\hat{c}_{jt} = (x_{jt}, w_{jt})\gamma + \omega_{jt},$$

    Store residuals, $\hat{\omega}_{jt}$, and $\hat{\theta}_3(\theta_2) = (\hat{\gamma})$.
  - **Criterion:**

    $$g(\theta) \equiv \begin{pmatrix} \frac{1}{\#} \sum_t \sum_j \hat{\xi}_{jt} Z^D_{jt} \\ \frac{1}{\#} \sum_t \sum_j \hat{\omega}_{jt} Z^S_{jt} \end{pmatrix}$$

Differentiation instruments

BLP Differentiation Instruments

Use $Z_{jt}^{D} \equiv \sum_{k \neq j} x_{kt}$: sum of characteristics of competing products.

Differentiation instruments

BLP Differentiation Instruments

Use $Z_{jt}^{D} \equiv \sum_{k \neq j} x_{kt}$: sum of characteristics of competing products.

- **Intuition:** Measures local "congestion" of product space
  - If no nearby competing products, firm has more "local monopoly" power.
- **Armstrong (2016):** For $T \to \infty$ asymptotics, then such instruments lose power.
  - $T \to \infty \Rightarrow \sum_{k \neq j} x_{kt}$ becomes uncorrelated with markups.

Optimal Instruments

- Chamberlain (1987): Derives "optimal instruments"
  - Reynaerts & Verboven (2014)
- Alternative: $\mathbb{E}(\xi_{jt}|Z_t) = 0$ motivates $\mathbb{E}\left[\xi_{jt}\varphi(Z_t)\right] = 0$ for some $\varphi(\cdot)$.
- Gandhi & Houde (2019): Use a measure of *distance* in characteristics space, $\|x_{jt} - x_{kt}\|$.

Outline

General Tricks

- **Fixed point:** Work directly on $\exp(\delta_{ij})$ rather than δ_{ij} to aovid many calls to $\exp(\cdot)$.
- **Vectorization vs. parallelization over t:** Consider whether to stack all $\{\boldsymbol{\delta}_t\}_{t=1}^{T}$ or to loop over them. (The problems are independent.)

Fixed Point Algorithms

Problem

$$\text{find } \delta_t \in \mathbb{R}^J \text{ s.t. } s_t(\delta) = \mathcal{S}_t.$$

Fixed Point Algorithms

Problem

$$\text{find } \delta_t \in \mathbb{R}^J \text{ s.t. } s_t(\delta) = \mathcal{S}_t.$$

	Jac.	Cost/it	Required it.	Rate
Iterating on Γ	÷	Minimal	Many	Linear
Newton	✓	Expensive	Fewer	Quadratic
Levenberg-Marquardt	✓	Expensive	Fewest	Quadratic
SQUAREM	÷	Small	Few	?

SQUAREM Update Step

- $\boldsymbol{\delta}_t^{[i+1]} := \boldsymbol{\delta}_t^{[i]} - 2\alpha^{[i]}\mathbf{r}^{[i]} + (\alpha^{[i]})^2\mathbf{v}^{[i]}$
- $\alpha^{[i]} = \frac{(\mathbf{v}^{[i]})'\mathbf{r}^{[i]}}{\mathbf{v}^{[i]\prime}\mathbf{v}^{[i]}}$
- $\mathbf{r}^{[i]} = \Gamma(\boldsymbol{\delta}_t^{[i]}) - \boldsymbol{\delta}_t^{[i]}$
- $\mathbf{v}^{[i]} = \Gamma(\Gamma(\boldsymbol{\delta}_t^{[i]})) - 2\Gamma(\boldsymbol{\delta}_t^{[i]}) + \boldsymbol{\delta}^{[i]}$

SQUAREM Intuition

SQUAREM Update Step

- $\boldsymbol{\delta}_t^{[i+1]} := \boldsymbol{\delta}_t^{[i]} - 2\alpha^{[i]}\mathbf{r}^{[i]} + (\alpha^{[i]})^2\mathbf{v}^{[i]}$
- $\alpha^{[i]} = \frac{(\mathbf{v}^{[i]})'\mathbf{r}^{[i]}}{\mathbf{v}^{[i]\prime}\mathbf{v}^{[i]}}$
- $\mathbf{r}^{[i]} = \Gamma(\boldsymbol{\delta}_t^{[i]}) - \boldsymbol{\delta}_t^{[i]}$
- $\mathbf{v}^{[i]} = \Gamma(\Gamma(\boldsymbol{\delta}_t^{[i]})) - 2\Gamma(\boldsymbol{\delta}_t^{[i]}) + \boldsymbol{\delta}^{[i]}$

SQUAREM Intuition

SQUAREM Update Step

- $\boldsymbol{\delta}_t^{[i+1]} := \boldsymbol{\delta}_t^{[i]} - 2\alpha^{[i]}\mathbf{r}^{[i]} + (\alpha^{[i]})^2\mathbf{v}^{[i]}$
- $\alpha^{[i]} = \frac{(\mathbf{v}^{[i]})'\mathbf{r}^{[i]}}{\mathbf{v}^{[i]\prime}\mathbf{v}^{[i]}}$
- $\mathbf{r}^{[i]} = \Gamma(\boldsymbol{\delta}_t^{[i]}) - \boldsymbol{\delta}_t^{[i]}$
  - $\cong \Gamma'(\cdot)$: jacobian
- $\mathbf{v}^{[i]} = \Gamma(\Gamma(\boldsymbol{\delta}_t^{[i]})) - 2\Gamma(\boldsymbol{\delta}_t^{[i]}) + \boldsymbol{\delta}^{[i]}$

SQUAREM Intuition

SQUAREM Update Step

- $\boldsymbol{\delta}_t^{[i+1]} := \boldsymbol{\delta}_t^{[i]} - 2\alpha^{[i]}\mathbf{r}^{[i]} + (\alpha^{[i]})^2\mathbf{v}^{[i]}$
- $\alpha^{[i]} = \frac{(\mathbf{v}^{[i]})'\mathbf{r}^{[i]}}{\mathbf{v}^{[i]\prime}\mathbf{v}^{[i]}}$
- $\mathbf{r}^{[i]} = \Gamma(\boldsymbol{\delta}_t^{[i]}) - \boldsymbol{\delta}_t^{[i]}$
    - $\cong \Gamma'(\cdot)$: jacobian
- $\mathbf{v}^{[i]} = \Gamma(\Gamma(\boldsymbol{\delta}_t^{[i]})) - 2\Gamma(\boldsymbol{\delta}_t^{[i]}) + \boldsymbol{\delta}^{[i]}$
    - Curvature: $f''(x) = \frac{f(x+2h) - 2f(x+h) - f(x)}{h^2}$

SQUAREM Intuition

SQUAREM Update Step

- $\boldsymbol{\delta}_t^{[i+1]} := \boldsymbol{\delta}_t^{[i]} - 2\alpha^{[i]}\mathbf{r}^{[i]} + (\alpha^{[i]})^2\mathbf{v}^{[i]}$

- $\alpha^{[i]} = \frac{(\mathbf{v}^{[i]})'\mathbf{r}^{[i]}}{\mathbf{v}^{[i]\prime}\mathbf{v}^{[i]}}$
  - Scalar: $[\Gamma''(\cdot)]^{-1}\Gamma'(\cdot)$, i.e. Newton step
- $\mathbf{r}^{[i]} = \Gamma(\boldsymbol{\delta}_t^{[i]}) - \boldsymbol{\delta}_t^{[i]}$
  - $\cong \Gamma'(\cdot)$: jacobian
- $\mathbf{v}^{[i]} = \Gamma(\Gamma(\boldsymbol{\delta}_t^{[i]})) - 2\Gamma(\boldsymbol{\delta}_t^{[i]}) + \boldsymbol{\delta}^{[i]}$
  - Curvature: $f''(x) = \frac{f(x+2h) - 2f(x+h) - f(x)}{h^2}$

SQUAREM Intuition

SQUAREM Update Step

- $\boldsymbol{\delta}_t^{[i+1]} := \boldsymbol{\delta}_t^{[i]} - 2\alpha^{[i]}\mathbf{r}^{[i]} + (\alpha^{[i]})^2\mathbf{v}^{[i]}$
  - Scalar: $\delta - 2\frac{r^2}{v} + \frac{r^2}{v^2}v = \delta - v^{-1}r^2$
- $\alpha^{[i]} = \frac{(\mathbf{v}^{[i]})'\mathbf{r}^{[i]}}{\mathbf{v}^{[i]\prime}\mathbf{v}^{[i]}}$
  - Scalar: $[\Gamma''(\cdot)]^{-1}\Gamma'(\cdot)$, i.e. Newton step
- $\mathbf{r}^{[i]} = \Gamma(\boldsymbol{\delta}_t^{[i]}) - \boldsymbol{\delta}_t^{[i]}$
  - $\cong \Gamma'(\cdot)$: jacobian
- $\mathbf{v}^{[i]} = \Gamma(\Gamma(\boldsymbol{\delta}_t^{[i]})) - 2\Gamma(\boldsymbol{\delta}_t^{[i]}) + \boldsymbol{\delta}^{[i]}$
  - Curvature: $f''(x) = \frac{f(x+2h) - 2f(x+h) - f(x)}{h^2}$

MPEC

- **Dube, Fox & Su (2012):** Suggest MPEC formulation of the BLP estimator.
  - Point out the issue with a nested tolerance.
  - Derive the relationship between the inner nested fixed point tolerance, and the bias in parameter estimates.

MPEC Formulation

$$\min_{\theta,\boldsymbol{\xi}} g(\xi)' W g(\xi)$$
$$\text{s.t.} \quad \mathbf{s}(\xi;\theta) = \boldsymbol{S}$$

- **Crucial:** Providing the sparsity structure (independence across markets wrt. ξ)

Bayesian Formulation

Yang, Chen & Allenby (2003); Jiang, Manchanda & Rossi (2009)

- **Cost:** Make a *functional form* assumption on ξ_{jt}
- **Reward:** Possible to simulate the entire system.
- **Benefit:** Posterior can be used to construct confidence intervals on non-linear functions such as price elasticities.
  - Normally: Delta Method or Bootstrapping required.

That's all