Minicurso IFB e Ibmec-SP: Métodos Empíricos Estruturais em Organização Industrial

XXVIII Encontro Brasileiro de Econometria

Salvador, BA, 6 e 7 de Dezembro de 2006

Leonardo Rezende

lrezende@econ.puc-rio.br

PUC-Rio e University of Illinois

Objetivos:

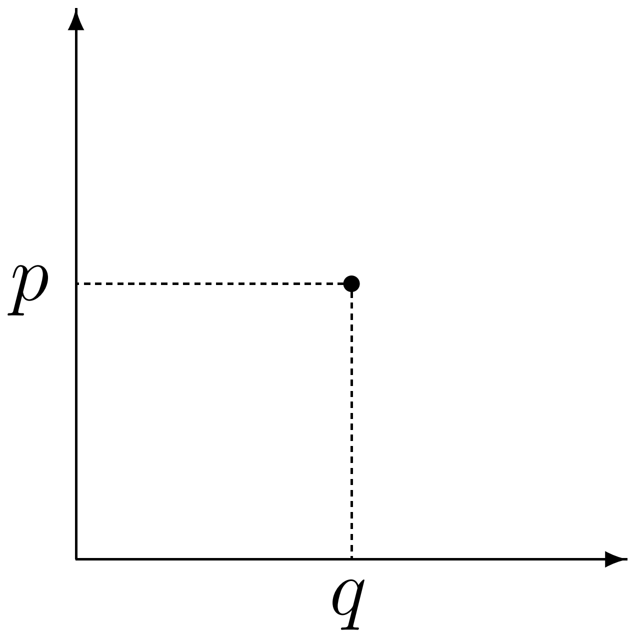
- Uma introdução à OI empírica estrutural
- 1ª Aula
  - Um exemplo: Bresnahan (1987)
- 2ª Aula
  - Discussão sobre como estimar demanda por bens heterogêneos: Logit, BLP, micro-BLP

Motivação

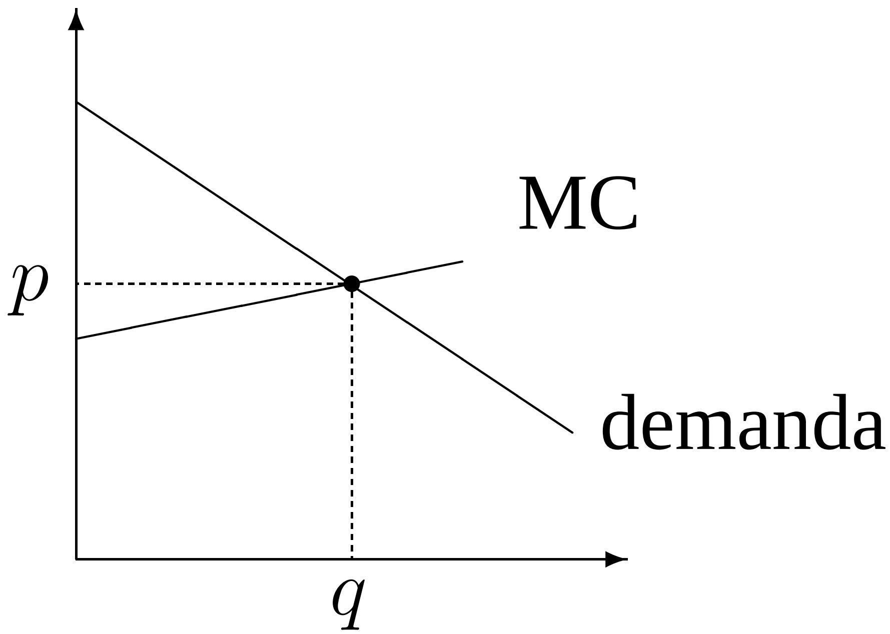
Como medir conduta anticompetitiva?
Não queremos/podemos usar:

- lucro contábil ($\neq$ lucro econômico)
- custo marginal (não observável)

O problema de identificação

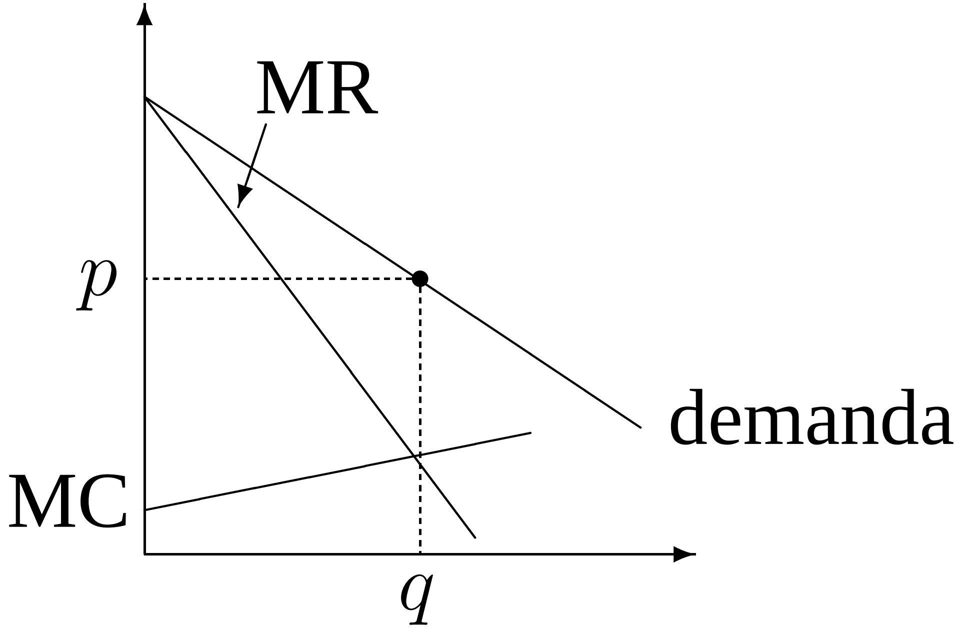


O problema de identificação



competição

O problema de identificação



monopólio

Especificações

Demanda (inversa):

$$p = D(q, Z, \delta, \epsilon_D)$$

- Z = regressores, δ = parâmetros, ϵ_D = resíduo

Custo Marginal:

$$MC(q_i, W, \Gamma, \epsilon_C)$$

- W = regressores, Γ = parâmetros, ϵ_C = resíduo

Sob competição:

$$p = D(q, Z, \delta, \epsilon_D)$$

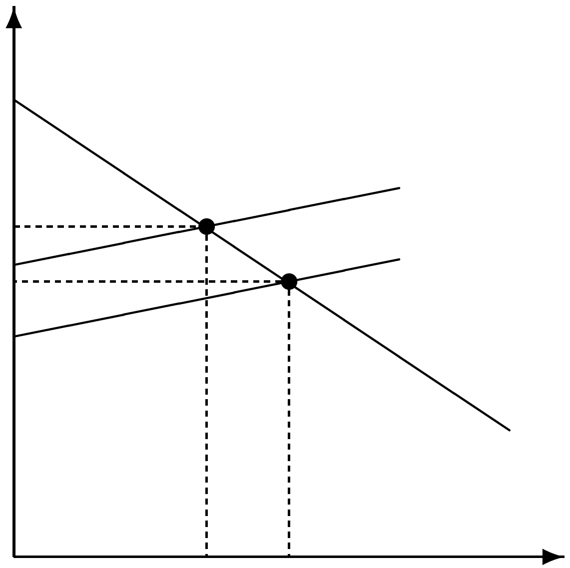
$$p = MC(q_i, W, \Gamma, \epsilon_C)$$

Monopólio:

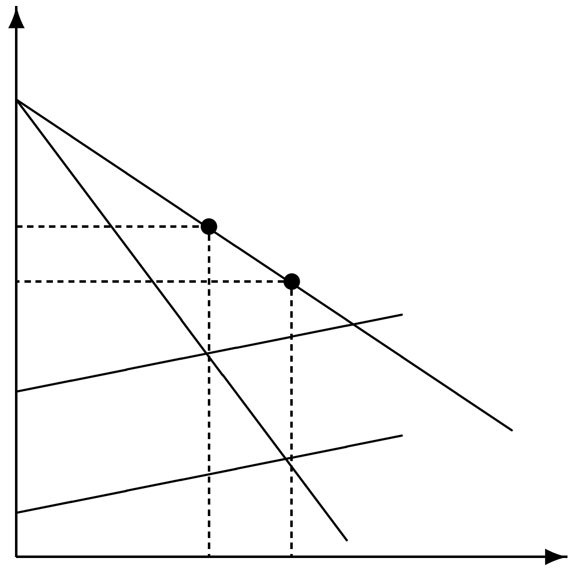
$$p = D(q, Z, \delta, \epsilon_D)$$

$$p - qD'(q, Z, \delta, \epsilon_D) = MR = MC(q_i, W, \Gamma, \epsilon_C)$$

Variação em W (em MC) identifica demanda:



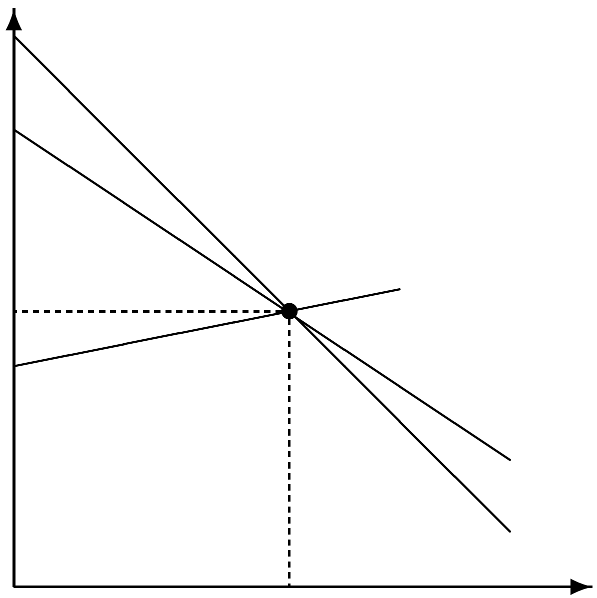
competição



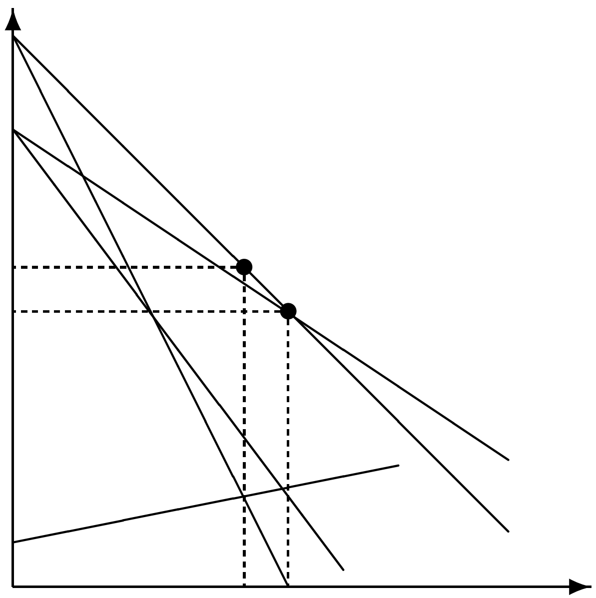
monopólio

... mas não conduta.

Para identificar conduta, precisamos variar D'!



competição



monopólio

Na prática, precisamos ver as mesmas firmas operando em mercados diferentes (= diferentes elasticidades)
Produtos diferenciados (e.g., automóveis):

- Compactos:
  - VW Gol, Chevrolet Celta, Fiat Palio, Renault Clio, Peugeot 206, Citroen C3
- Compactos “arrendondados”:
  - VW New Beetle

Bresnahan (1987)

"Competition and Collusion in the American Automobile Industry: the 1955 Price War" *Journal of Industrial Economics* XXXV n.4, June 1987.

Sobre o "puzzle de 1955":

"(I) would flunk any econometrics paper that claimed to provide an explanation of 1955 auto sales"

— Paul Samuelson

Bresnahan: Mudança de Conduta:

- Competição em 1955
- Conluio em 1954 e 1956

Demanda por automóveis

- bem heterogêneo: dezenas de modelos
- Escolha discreta: demanda individual = 0 ou 1 (ou 2?...)

Bresnahan: Diferenciação Vertical:

- Modelo i tem *qualidade* x_i e preço p_i.

x_i:

- É o mesmo para todos os consumidores
- (não diretamente observado por nós)

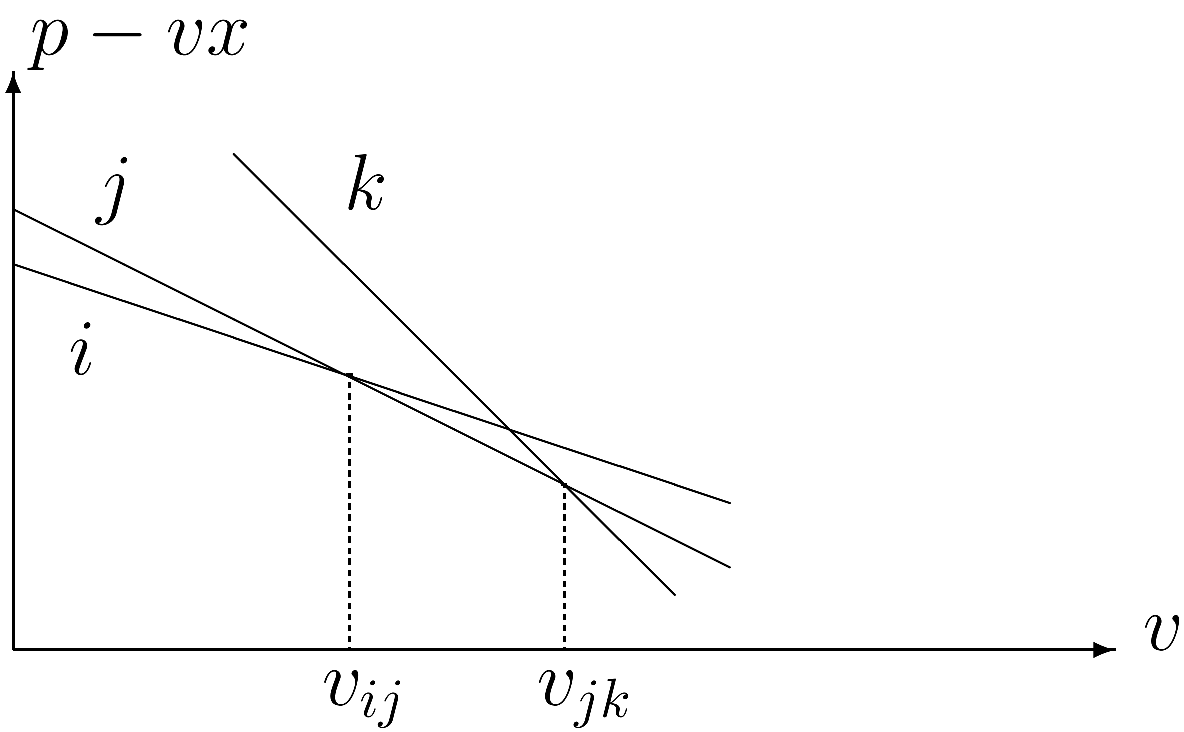
Consumidores têm diferentes propensões marginais a pagar por qualidade:

$$v \sim U[0, v_{max}], \text{densidade } \delta$$

$$U(\text{carro com } (x_i, p_i)) = vx_i + Y - p_i$$

$$U(\text{sem carro}) = v\gamma + Y - E$$

Consumidor escolhe carro que minimize “preço líquido” $p - vx$:



Calculando a Demanda

$$p_i - x_i v_{ij} = p_j - x_j v_{ij} \Rightarrow v_{ij} = \frac{p_j - p_i}{x_j - x_i}$$

Demanda por carro j:

$$q_j = \delta [v_{jk} - v_{ij}] = \delta \left[\frac{p_k - p_j}{x_k - x_j} - \frac{p_j - p_i}{x_j - x_i} \right]$$

$$q_1 = \delta \left[\frac{p_2 - p_1}{x_2 - x_1} - \frac{p_1 - E}{x_1 - \gamma} \right]$$

$$q_n = \delta \left[v_{max} - \frac{p_n - p_{n-1}}{x_n - x_{n-1}} \right]$$

Obs.:

$$D' = -\delta \left[\frac{-1}{x_k - x_j} - \frac{1}{x_j - x_i} \right]$$

... varia com $|x_i - x_j|$: Elasticidade maior se há carros de qualidade similar.

Lado da firma

Custo:

$$C(x, q) = A(x) + mc(x)q;$$

- $A(x)$ irrelevante
- $mc(x) = \mu e^x$

Lucro (do modelo j):

$$\pi_j = (p_j - mc(x_j))q(p_i, p_j, p_k, x_i, x_j, x_k) - A(x_j)$$

Lucro da firma f:

$$\Pi_f = \sum_{j \text{ vendido por } f} \pi_j$$

Preços de Equilíbrio e Conduta

- x_i exógeno.
- p_i determinado em equilíbrio.

Dois tipos de conduta possíveis:

- "Competição": Bertrand-Nash
- Conluio: Monopólio

Bertrand Nash

Se firma f produz j e não produz i e k:

$$q_j(p,x) + (p_j - mc(x_j))\frac{\partial}{\partial p_j} q_j(x) = 0$$

Se produz k também:

$$q_j + (p_j - mc_j)\frac{\partial}{\partial p_j} q_j + (p_k - mc_k)\frac{\partial}{\partial p_j} q_k = 0$$

etc.
Dado $x = (x_1, \ldots, x_n)$, temos sistema de n equações em $p = (p_1, \ldots, p_n)$.

Conluio

p de conluio: solução para

$$q_j + (p_j - mc_j)\frac{\partial}{\partial p_j}q_j + (p_k - mc_k)\frac{\partial}{\partial p_j}q_k + (p_i - mc_i)\frac{\partial}{\partial p_j}q_i = 0$$

para todos os carros.

(Como se firmas fossem uma só.)

Especificação Econométrica

Dados:

- p_i (list price)
- q_i (total de carros i produzidos nos EUA num ano)
- z_i: vetor de características do modelo i
  - peso, comprimento, potência, # cilindros, se conversível

$$x_i = \sqrt{\beta_0 + \sum_t z_{it}\beta_t}$$

Definição de modelos: não trivial.

Algoritmo

Parâmetros: $\beta, \mu, \delta, v_{max}, \gamma, E(= mc(\gamma))$.

1. Com chute de β, calcular $x_1, \ldots, x_n$.
2. Ordenar modelos de acordo com x_i.
3. Com chute de $\mu, \delta, v_{max}, \gamma$, calcular $p_1^*, \ldots, p_n^*, q_1^*, \ldots, q_n^*$ de equilíbrio.
4. Calcular verossimilhança, supondo que:

$$p_i = p_i^* + \epsilon_i^p, \qquad \epsilon_i^p \sim N(0, \sigma_p^2)$$

$$q_i = q_i^* + \epsilon_i^q, \qquad \epsilon_i^q \sim N(0, k_i \sigma_q^2)$$

Obs.: ϵs não são erros estruturais; erros de “forma reduzida”, ou de mensuração.

Estimação

Bresnahan estima o modelo separadamente para:

- 1954, 1955, 1956
- Competição e Conluio
- (outras duas especificações: hedônica e "produtos")

Dois testes:

- formal: Teste de Cox (porque $H_0 \not\subset H_1$).
- informal: Comparação de Coeficientes

Teste de Cox

Idéia: Sob H_0,

$$\sqrt{n}\left(\frac{1}{n}(L_0 - L_1) - E_0[L_0 - L_1]\right) \to N$$

Mais precisamente, sejam

- s^2_{0p}, s^2_{0q} variâncias estimadas sob H_0;
- s^2_{1p}, s^2_{1q} variâncias estimadas sob H_1;
- ϕ^2_p, ϕ^2_q variâncias estimadas sob H_1 usando p^*, q^* de H_0 como dados;

Então

$$T = \frac{n_p}{2}\log\left(\frac{s^2_{1p}}{s^2_{0p} + \phi^2_p}\right) + \frac{n_q}{2}\log\left(\frac{s^2_{1q}}{s^2_{0q} + \phi^2_q}\right)$$

Características de OI empírica estrutural

- Foco: em uma indústria, ou uma firma, ou uma pessoa(!)
- Sob medida: metodologia específica para cada mercado.
- Teoria Econômica levada a sério: agentes e econometrista usam o mesmo modelo.
- Computacionalmente complexo.
- Formas funcionais restritivas.
- Propriedades formais dos estimadores nem sempre claras.
- Resultados nem sempre convincentes (Bresnahan é exceção!)

Motivação para próxima aula

- Aprendemos que a hipótese crítica é sobre elasticidades-preço.
- Modelo de diferenciação vertical gera variação (via xs), mas…
- elasticidade preço cruzada entre j e $n-3$ não vizinhos tem que ser 0!!
- Como ficam VW Gol, Fiat Palio, Renault Clio, Peugeot 206, Citroen C3?

Na aula dois, veremos outros sistemas de demanda para bens diferenciados mais flexíveis.

Segunda Aula

Queremos métodos para estimar demanda por bens diferenciados que:

- venham da agregação de decisões individuais;
- sejam estimáveis, mesmo com muitos bens;
- gerem elasticidades cruzadas mais realistas.

Discutiremos modelos de escolha discreta:

- Logit
- Nested Logit
- BLP
- micro-BLP

Precisamos desses modelos?

Por que não estimar simplesmente sistema de demanda

$$\begin{aligned}\log q_1 &= X_1\beta + \eta_{11}\log p_1 + \ldots + \eta_{1n}\log p_n + u_1\\ &\vdots\\ \log q_n &= X_n\beta + \eta_{n1}\log p_1 + \ldots + \eta_{nn}\log p_n + u_n?\end{aligned}$$

Inviável para n grande:

- n^2 elasticidades;
- n intrumentos;
- às vezes, apenas n observações.

Logit

Logit multinomial:

- Modelo de escolha discreta mais popular
- deu o Nobel de 2000 a McFadden

Consumidor i tem $n+1$ opções:

- n bens + “outside good” (bem 0).

Consumidor resolve:

$$\max_{j \in \{0,\dots,n\}} u_{ij}$$

onde

$$u_{ij} = \delta_j + \epsilon_{ij} = X_j\beta - \alpha p_j + \epsilon_{ij}$$

- δ_j: componente sistemático
- ϵ_{ij}: componente idiossincrático (=resíduo)

Obs. $u_{ij} \neq$ utilidade, utilidade indireta.

Consumidor compra j se

$$\delta_j + \epsilon_{ij} \geq \delta_k + \epsilon_{ik}, \forall k.$$

- M consumidores nesse mercado ($M >> n$).
- ϵ_{ij} i.i.d.

Pela LGN,

$$\begin{aligned} s_j &= \frac{q_j}{M} = \frac{1}{M} \sum_i \mathbb{I}\{\delta_j + \epsilon_{ij} \geq \delta_k + \epsilon_{ik}, \forall k\} \\ &\rightarrow \Pr(\delta_j + \epsilon_{ij} \geq \delta_k + \epsilon_{ik}, \forall k) \end{aligned}$$

Truque: $\epsilon \sim F(t) = e^{-e^{-t}}$
(Distribuição valor extremo tipo 1)

$$\begin{aligned}
s_j &= \Pr(\epsilon_{ik} \leq \epsilon_{ij} + \delta_j - \delta_k, \forall k) = \int \left(\prod_{k \neq j} e^{-e^{-(t+\delta_j-\delta_k)}} \right) e^{-t} e^{e^{-t}} dt \\
&= \int \left(\prod_k e^{-e^{-(t+\delta_j-\delta_k)}} \right) e^{-t} dt = \int \left(e^{-e^{-t} \sum_k e^{-(\delta_j-\delta_k)}} \right) e^{-t} dt \\
&= \int_0^\infty e^{-x \sum_k e^{-(\delta_j-\delta_k)}} dx = \left. \frac{e^{-x \sum_k e^{-(\delta_j-\delta_k)}}}{-\sum_k e^{-(\delta_j-\delta_k)}} \right|_0^\infty \\
&= \frac{1}{\sum_k e^{-(\delta_j-\delta_k)}} = \frac{e^{\delta_j}}{\sum_k e^{\delta_k}}.
\end{aligned}$$

$$s_j = \frac{e^{\delta_j}}{\sum_k e^{\delta_k}} \Rightarrow \frac{s_j}{s_k} = \frac{e^{\delta_j}}{e^{\delta_k}}$$

Normalização: $\delta_0 = 0$.

$$\frac{s_j}{s_0} = e^{\delta_j}$$

$$\log s_j - \log s_0 = \delta_j = X_j\beta - \alpha p_j$$

β, α podem ser estimados por OLS.

“Red Bus, Blue Bus”

Conta depende da definição de produtos:

- Se dividirmos j em j' e j'' (com $\delta_j = \delta_{j'} = \delta_{j''}$),

$$s_j < s_{j'} + s_{j''} (\simeq 2 s_j)$$

Elasticidades

- p_k afeta δ_k que afeta $s_j = \frac{e^{\delta_j}}{\sum_k e^{\delta_k}}$, para todo j.
- Porém $\frac{s_j}{s'_j} = \frac{e^{\delta_j}}{e^{\delta_{j'}}}$ *não* muda.
  - Uma promoção no Clio afeta na mesma proporção vendas de Peugeots 206 e Ferraris.

IIA: Independence of Irrelevant Alternatives

Nested Logit

Aninhar escolha em grupos:

- Compactos, SUVs, Luxo, etc.

Utilidade para bem j do grupo g:

$$u_{ij} = \delta_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij},$$

com ϵ_{ij} e $\zeta_{ig} + (1-\sigma)\epsilon_{ij}$ ~e.v.t.1.

- σ mede importância do grupo.

No Nested Logit,

$$\log s_j - \sigma \log(s_{j|g}) - \log s_0 = \delta_j$$

onde $s_{j|g}$ é o market share de j dentro de g.

BLP

BLP: Berry, Levinsohn e Pakes (1995): “Automobile Prices in Market Equilibrium”, *Econometrica* 63, n.4.

- Principal Referência na Literatura
- Como Bresnahan,
  - sobre mercado de automóveis nos EUA, 1971-1990
  - dados anuais de p_j, q_j e z_j agregado por modelo
  - Estima demanda e oferta, preços de oligopólio
- diferente de Bresnahan,
  - especificação de demanda muito mais geral.

A demanda BLP

Ponto de partida:

$$u_{ij} = X_j\tilde{\beta}_i - \alpha p_j + \xi_j + \epsilon_{ij}$$

Duas extensões cruciais:

- ξ_j: característica não observável de j
- $\tilde{\beta}_i$: heterogeneidade de preferências

O ξ_j

Características não observadas do bem:

- Elementos subjetivos (beleza do carro); erros de medição de z_j; quaiquer outros efeitos que fçam j desejável para todos os consumidores.

... estão em ξ_j.

Considere modelo $u_{ij} = X_j\beta - \alpha p_j + \chi_j + \epsilon_{ij}$.
Derivação do Logit ainda vale, com

$$\delta_j = X_j\beta - \alpha p_j + \chi_j.$$

Obtemos:

$$\log s_j - \log s_0 = \delta_j = X_j\beta - \alpha p_j + \xi_j$$

ξ_j está (é) resíduo da regressão original.
Problema:

- p_j é endógeno (firmas conhecem ξ_j).

Precisaremos de instrumentos.

Os $\tilde{\beta}_i$

Consumidores podem ter diferentes preferências sobre características:

- tamanho pode ser bom para famílias grandes, ruim para solteiros.

Supomos que coeficiente para característica t é

$$\tilde{\beta}_{ti} = \beta_t + \zeta_{ti}, \qquad \zeta_{ti} \sim N(0, \sigma_t^2)$$

então

$$u_{ij} = [X_j\beta - \alpha p_j + \xi_j] + [\sum_t x_{jt}\zeta_{ti} + \epsilon_{ij}] = \delta_j + \nu_{ij}.$$

$\nu_{ij} \not\sim$ e.v.t.1. Fórmulas do logit não são válidas.

$$s_j = S_j(\delta) = \int \left(e^{\delta_j} / \sum e_k^{\delta} \right) dP_\sigma(\zeta)$$

Algoritmo tem que resolver sistema para obter:

$$S^{-1}(s) = \delta_j = X_j\beta - \alpha p_j + \xi_j$$

(ver Berry 1994)

Renda

$y_i =$ Renda de i.

$$u_{ij} = X_j \tilde{\beta}_i - \alpha \log(y_i - p_j) + \xi_j + \epsilon_{ij}$$

(Demanda de ricos mais inelástica).
Problema:

- Não há dados de compras por nível de renda.

Solução (parcial): tratar y_i como um ν:

$$y_i \sim \text{lognormal},$$

Com parâmetros estimados com CPS de cada ano.

Oferta

$$\log(mc_j) = W_j\gamma + \omega_j$$

- W_j análogo (=) a X_j
- ω_j análogo a ξ_j

Lucro da firma é

$$\Pi_f = \sum_{j \text{ de } f} (p_j - mc_j) M s_j(p)$$

CPO

$$s_j + \sum_{r \text{ de } f} (p_r - mc_r) \frac{\partial}{\partial p_j} s_r = 0$$

Definindo matriz $n \times n$: Δ

$$\Delta_{jr} = -\frac{\partial}{\partial p_j} s_r \text{ se } r, j \text{ da mesma firma ou } 0$$

CPO fica

$$s - \Delta[p - mc] = 0$$

$$mc = p - \Delta^{-1} s$$

$$\log(p - \Delta^{-1} s) = W\gamma + \omega$$

Instrumentos

- ξ_j, ω_j afetam decisões das firmas, como p_j.
- Características de produtos x_j, w_j supostas exógenas.

Idéia:

- usar x_r:
  - em oligopólio, x_r afeta mark-up de $j \Rightarrow p_j$
  - *suposto* exógeno.

Algoritmo

Dado "chute" de parâmetros,

1. simule $\nu \sim P_\sigma$ para poder calcular $S_j(\delta) = \int \left(e_j^\delta\right) / \sum e_k^\delta \right) dP_\sigma$
2. itere
$$\delta^{t+1} = \delta^t + \log(s) - \log(S(\delta^t))$$
para obter δ.
3. $\xi = \delta - x\beta$.
4. calcule $\Delta(= -\frac{\partial}{\partial p_j} s_r$ e zeros).
5. $\omega = \log(p - \Delta^{-1} s) - W\gamma$.
6. GMM em $E((\xi, \omega)|z) = 0$.

micro-BLP

Berry, Levinsohn and Pakes, “Differentiated Products Demand System from a Combination of Micro and Macro Data: The New Car Market”, *Journal of Political Economy*, 112, 1, Feb. 2004.

Como combinar dados agregados (ao nível do produto) com microdados (ao nível do consumidor). Microdados:

- características do consumidor $\times$ compra
- inclui dados sobre segunda opção

- j: produto
- k: características dos produtos
- i: consumidor
- r: características do consumidor

$$u_{ij} = \sum_k x_{jk}\tilde{\beta}_{ik} + \xi_j + \epsilon_{ij}$$

agora,

$$\tilde{\beta}_{ik} = \bar{\beta}_k + \sum_r z_{ir}\beta^o_{kr} + \beta^u_k \nu_{ik}$$

(Não) lidando com ξ

Podemos escrever

$$u_{ij} = \delta_j + \sum_{k,r} x_{jk} z_{ir} \beta^o_{kr} + \sum_k x_{jk} \beta^u_k \nu_{ik} + \epsilon_{ij}$$

onde $\delta_j = \sum_k x_{jk} \bar{\beta}_k + \xi_j$.
Considere δ_j como parâmetro (efeito fixo):

- Podemos estimar β^o, β^u sem fazer nenhuma hipótese sobre ξ.
- Mas não podemos estimar $\bar{\beta}$.

Precisamos de $\bar{\beta}_k$?

Sim: $\bar{\beta}_k$ de p (=$-\alpha$) necessário para elasticidades.
Mas *padrão* de substituições determinado por β^o, β^u.
Elasticidade é composição de dois efeitos:

1. Se p_j sobe, δ_j cai (medido por $\bar{\beta}$)
2. Se δ_j cai, $s_{j'}$ sobe (determinado por β^o, β^u)

No Logit, $\beta^o, \beta^u = 0$.

Método: GMM em

1. s_j previsto no modelo $=$ nos dados

   $\longrightarrow$ determina δ

2. Cov$(x$ do carro comprado, $z)$ previsto no modelo $=$ nos dados

   $\longrightarrow$ para estimar β^o

3. Cov$(x$ do carro comprado, x da segunda escolha) previsto no modelo $=$ nos dados

   $\longrightarrow$ para estimar β^u

Se i compra j (com x_k alto) de forma inesperada (dado z_i), ou $\epsilon_{ij} >> 0$ ou $\nu_{ik} >> 0$. Se $\nu_{ik} >> 0$, segunda escolha será carro com x_k alto.