Notas de Aula:

Demanda por Bens Diferenciados

Victor Gomes
Universidade de Brasília
victorgomes@unb.br

7 de novembro de 2010

1 Modelo

Seguimos aqui o artigo de Nevo (2000).
mercado: $t = 1, ..., T$
consumidores por mercado: $i = 1, ..., I_t$
informações por mercado: (i) quantidades, (ii) preço médio (p_{it}), (iii) características dos produtos (J produtos)

Função de utilidade indireta para o consumidor i, consumindo o produto j no mercado t.

$$u(x_{jt}, \epsilon_{jt}, p_{jt}, \tau_i; \theta) \tag{1}$$

tal que:
x_{jt}: características observadas do produto j no mercado t
ϵ_{jt}: características não-observadas pelo pesquisador/economista
p_{jt}: preço médio do produto j no mercado t
τ_i: características dos indivíduos
θ: parâmetros desconhecidos

Especificação da função utilidade

$$u_{ijt} = \alpha_i(y_i - p_{jt}) + x_{jt}\beta_i + \epsilon_{jt} + \epsilon_{ijt} \tag{2}$$

- $i = 1, ..., I_t$ (consumidores por mercado)
- $j = 1, ..., J$ (produtos)
- $t = 1, ..., T$ (mercados)
- x_{jt} é um vetor de dimensão K de características observáveis do produto j.

- ϵ_{ijt} termo estocástico, $\mathrm{E}[\epsilon] = 0$.
- α_1 utilidade marginal do consumidor da renda.
- β_i vetor de gostos dos indivíduos

Obs: as características variam com o tipo de produto estudado.

Características não-observadas podem incluir o impacto de atividade promocional, fatores não-quantificáveis (marca) ou choques sistemáticos de demanda. Ex:

$$\epsilon_{jt} = \underbrace{\epsilon_j}_{\text{dummy marca}} + \underbrace{\epsilon_t}_{\text{dummy mercado}} + \Delta\epsilon_{jt} \tag{3}$$

Três observações a respeito da função (2).

1. função utilidade
   - BLP – Cobb-Douglas
   - Nevo – quasi-linear (sem efeito-riqueza)
2. características não-observáveis são idênticas para todos os consumidores (α_i varia entre indivíduos – consistente com literatura de diferenciação vertical dos produtos)
3. todos os consumidores se deparam com as mesmas características dos produtos – o mesmo preço é oferecido aos consumidores.

- As preferências variam como função das características individuais τ_i.
- Duas características (tipos):
  1. Observada (D_i): renda, idade, tamanho família, educação, etc.
  2. Não-observada (v_i): sem informação. ex: animal estimação, religião, etc.

Formalmente:

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i \tag{4}$$

tal que

$$v_i \sim P_v^*(v)$$

$$D_i \sim \hat{P}_D^*(D)$$

são independentes. $P(\ )$ é uma distribuição paramétrica ou não-paramétrica.

Π : (K + 1) $\times$ d : matriz dos coeficientes que varia com D_i.

Σ : (K + 1) $\times$ (K + 1) parâmetros : matriz dos coeficientes de v_i.

(Implicação: taste parameter varia com D_i; + informação para análise).

Outside good : os consumidores podem decidir por não comprar nenhuma marca (produto).

$$u_{i0t} = \alpha_i y_i + \underbrace{\epsilon_{0t} + \pi_0 D_i + \delta_0 v_{i0}}_{\text{zero}} + \epsilon_{i0t} \tag{5}$$

$$u_{i0t} \simeq 0$$

$$\alpha_i y_i \simeq 0$$

- - - - X - - - -

Parâmetros desconhecidos $\theta = (\theta_1, \theta_2)$.

$\theta_1 = (\alpha, \beta)$ – parâmetros lineares

$\theta_2 = (\Pi, \Sigma)$ – parâmetros não-lineares

Combinando as equações (2) e (4) temos:

$$u_{ijt} = \alpha_i y_i - \alpha_i p_{jt} + x_{jt}\beta_i + \epsilon_{jt} + [-p_{jt}, x_{jt}](\Pi D_i + \Sigma v_i) + \epsilon_{ijt} \tag{6}$$

$$u_{ijt} = \alpha_i y_i + \delta_{jt}\left(p_{jt}, x_{jt}, \epsilon_{jt} : \theta_1\right) + \mu_{ijt}\left(p_{jt}, x_{jt}, D_i, v_i; \theta_2\right) + \epsilon_{ijt} \tag{7}$$

tal que:

$$\delta_{jt}\left(p_{jt}, x_{jt}, \epsilon_{jt} : \theta_1\right) = -\alpha_i p_{jt} + x_{jt}\beta_i + \epsilon_{jt} \tag{8}$$

$$\mu_{ijt}\left(p_{jt}, x_{jt}, D_i, v_i; \theta_2\right) = \underbrace{[-p_{jt}, x_{jt}]}_{[1\times(K+1)]}(\Pi D_i + \Sigma v_i) \tag{9}$$

$\delta_{jt}(\cdot)$: utilidade média – comum a todos os consumidores.

$\mu_{ijt}(\cdot) + \epsilon_{ijt}$: efeitos aleatórios – desvios média-zero heterocedasticos da utilidade média.

Hipótese consumidores compram uma unidade do bem que fornece maior utilidade.

- Como o indivíduo é definido como um vetor demográfico e choques específicos dos produtos $(D_i, v_i, \epsilon_{i0t}, ..., \epsilon_{ilt})$ isto implicitamente define o conjunto de atributos individuais que levam à escolha do bem j.
- Formalmente:

$$A_{jt}(x_{\cdot t}, p_{\cdot t}, \delta_{\cdot t}; \theta_2) = \{(D_i, v_i, \epsilon_{i0t}, ..., \epsilon_{ilt}) \mid u_{ijt} \leq u_{ilt} \forall\ l = 0, 1, ..., J\}, \quad (10)$$

Para todas as marcas:

- $x_{\cdot t} = (x_{lt}, ..., x_{Jt})$ (características)
- $p_{\cdot t} = (p_{lt}, ..., p_{Jt})$ (preços médios)
- $\delta_{\cdot t} = (\delta_{lt}, ..., \delta_{Jt})$ (utilidade média)

Market Share

$$s_{jt}(x_{\cdot t}, p_{\cdot t}, \delta_{\cdot t}; \theta_2) = \int_{A_{jt}} \underbrace{P^*(D, v, \epsilon)}_{\text{f. dens. pop.}} \quad (11)$$

$$= \int_{A_{jt}} \underbrace{dP^*(\epsilon \mid D, v) dP^*(v \mid D) dP^*_D(D)}_{\text{pela regra de Bayes}} \quad (12)$$

Dada a hipótese de independência (13)

$$= \int_{A_{jt}} dP^*(\epsilon) dP^*_v(v) d\hat{P}^*_D(D) \quad (14)$$

Dadas as hipóteses sobre a distribuição dos atributos individuais não-observáveis podemos computar (11). Int: analiticamente ou numericamente.

Hipóteses Distribucionais

- Importantes para elasticidade-preço da demanda.
- se modelarmos os choques de forma aditiva e separável, então

$$\theta_2 = 0, \text{ ou } \beta_i = \beta \text{ e } \alpha_i = \alpha.$$

Neste caso estamos no modelo logit (agregado). Isto implica em

$$s_{jt} = \frac{\exp\{x_{jt}\beta - \alpha p_{jt} + \epsilon_{jt}\}}{1 + \sum_{k=1}^{J} \exp\{x_{kt}\beta - \alpha p_{kt} + \epsilon_{kt}\}} \quad (15)$$

- A elasticidade-preço da demanda sugerido pelos market shares em (15) são:

$$\eta_{jkt} = \frac{\partial s_{jt} p_{kt}}{\partial p_{kt} s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) d\hat{P}_D^*(D) dP_v^*(v) \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} d\hat{P}_D^*(D) dP_v^*(v) \end{cases} \tag{16}$$

  tal que

$$s_{ijt} = \frac{\exp\{\delta_{jt} + \mu_{ijt}\}}{1 + \sum_{k=1}^{K} \exp\{\delta_{kt} + \mu_{ikt}\}}$$

  é a probabilidade do indivíduo i comprar o produto j.

2 Estimação

2.1 Dados

Dados usados por Nevo.

- global: características da marca (produto)
- por mercado: market share e preço médio
- distribuição de informações demográficas ($\hat{P}_D^*$)

Observações:

- definir o market share em termos de bem-real.
- definir o tamanha do mercado. Importante! Definição correta do *outside good*.
- marca – dummy por marca.
  - # mercados > # marcas

2.2 Identificação

Algoritmo de estimação.

Estimador GMM:

$$\min_{\theta} ||s(x, p, \delta(x, p, \epsilon; \theta_1); \theta_2) - S|| \tag{17}$$

- $s(\cdot)$: market share do modelo
- S: market share observado
- Problema: os parâmetros não são lineares

Berry's solution Faça $z = [z_1, ..., z_M]$ ser um conjunto de instrumentos tal que

$$\mathrm{E}[z_m \omega(\theta^*)] = 0 \ , \ m = 1, ..., M$$

tal que ω é uma função dos parâmetros do modelo e um termo de erro a ser definido e θ^* representa os valores verdadeiros de θ.

O estimador GMM é o seguinte:

$$\hat{\theta} = \arg\min_{\theta} \left\{ \omega(\theta)' z \Phi^{-1} z' \omega(\theta) \right\} \tag{18}$$

- Φ é uma estimativa consistente de $\mathrm{E}[z'\omega\omega' z]$.
- Φ^{-1} define a métrica para o quão próximo de zero estamos. / inverso da matriz Var-Cov – entao damos menos peso para as eqs com maior variância.

Modelando o termo de erro ϵ_{jt} é estrutural – não é a diferença entre o observado e o previsto.

Precisamos determinar o termo de erro. De (7) temos:

$$u_{ijt} = \alpha_i y_i + \underbrace{\delta_{jt}(\cdot)}_{\epsilon_{jt}\uparrow} + \mu_{ijt}(\cdot) + \epsilon_{ijt} \tag{19}$$

Aqui $\delta_{jt}(\cdot)$ é uma função linear de ϵ_{jt}. Podemos expressar o termo de erro como uma função linear destes parâmetros.

Para cada mercado:

$$s(\delta_{\cdot t}; \theta_1) = S_{\cdot t} \tag{20}$$

Solucionando

$$\begin{aligned} s_{jt} &= \frac{1}{ns} \sum_{i=1}^{ns} s_{jti} \\ &= \frac{1}{ns} \sum^{n} s_{i=1} \frac{\exp\left\{ \delta_{jt} + \sum_{k=1}^{K} x_{jt}^k (\sigma_k v_i^k + \Pi_{k1} D_{i1} + ... + \Pi_{kd} D_{id} \right\}}{1 + \sum_{m=1}^{J} \exp\left\{ \delta_{mt} + \sum_{k=1}^{K} x_{mt}^k (\sigma_k v_i^k + \Pi_{k1} D_{i1} + ... + \Pi_{kd} D_{id} \right\}} \end{aligned}$$

tal que

$$(v_i^1, ..., v_i^k) \ \text{ amostra de } \hat{P}_v^*(v)$$

$$(D_{i1}, ..., D_{id}) \ \text{ amostra de } P_D^*(D)$$

$i = 1, ..., ns$, e $k = 1, ..., K$

- x_{jt}^k, $k = 1, ..., K$ são variáveis com coeficientes de inclinação aleatórios.

- $P^*_\epsilon(\epsilon)$ segue uma distribuição de valor extremo.

- Usando o market share calculado podemos inverter o sistema de equações (20).

O sistema pode ser usado utilizando o 'mapeamento' de contração:

$$\delta_t^{h+1} = \delta_t^h + \ln S_{\cdot t} - \ln s(p_{\cdot t}, x_{\cdot t}, delta^h_{\cdot t}, P_{ns}; \theta_2) \tag{21}$$

tal que $t = 1, ..., T$ e $h = 0, ..., H$.

H é o menor inteiro tal que

$$||\delta^H_{\cdot t} - \delta^{H-1}_{\cdot t}||$$

para um nível determinado de tolerância.

$\delta^H_{\cdot t}$ é uma aproximação de $\delta_{\cdot t}$.

Uma vez que a inversão tenha sido calculada, o termo de erro é definido como:

$$\omega_{jt} = \delta_{jt}(s_t; \theta_2) - (x_{jt}\beta + \alpha p_j t) \equiv \epsilon_{jt} \tag{22}$$

Comentário (intuição): para os valores dados do parâmetro não-linear θ_2, nós solucionamos para os níveis de utilidade média $\delta_{\cdot t}(\cdot)$, que determina que os market shares previstos são iguais aos market shares observados.

O resíduo ϵ_{jt} é a diferença entre essa valoração e a prevista pelos parâmetros lineares α e β.

2.3 Instrumentos

Necessidade: um conjunto de variáveis instrumentais exógenas.

- Obs
  - variáveis que deslocam os custos mas que não são correlacionadas com a demanda.
  - cost shiffter within brand?!

- BLP:
  - características observadas dos produtos
  - soma dos valores das mesmas marcas
  - E$(\epsilon \mid x) = 0$ características de outros produtos oferecidos pela firma.
  - soma dos valores das mesmas características produzidos por outras firmas.

- Nevo – Problemas/dúvidas sobre a validade dos instrumentos de BLP.

- as características devem (ou deveriam) ser determinadas de uma expectativa de ϵ.
- Se as características mudam muito por que não são correlacionadas com ϵ? Se não mudam por que estimativa de FE?
- Estrutura de painel (Nevo). Utiliza o preço de marcas de outras cidade como VI.

$$p_{jct} = mc_{jct} + \underbrace{mup_{jct}}_{\text{mark-up}} = mc_{jt} + \Delta mc_{jct} + mup_c + mup_t + \Delta mup_{jct} \quad (23)$$

- Preços são correlacionados por meio de choques comuns ao mc (mc_{jt}; assuma que os choques de demanda (ϵ_{jct}) são independentes, condicional uma vez que são controlados por marca, cidade (c) e efeitos fixos de tempo.
- Não tão popular. Requer mais dados. Não-intuitivo. Várias razões teóricas para não funcionar. (promoção nacional / propaganda; mudança na preferência pela marca).

2.4 Dummy para Marca

- Berry'94 argumenta que a metodologia de estimação é primariamente construída para instrumentalizar a correlação entre preços e ϵ_{jt}.
- Usando a dummy, ϵ_{jt}, não é apenas as característica não-observáveis, ela se torna desvios da média não-observada especifica aos vários mercados t.
- Esta variância surge quando tratamos de mais de um mercado especificamente.
- Os coeficientes de preferência βs podem não ser identificados. βs podem ser identificados usando o procedimento de distância mínima de Chamberlain'82.

$$d = (d_1, ..., d_j)'$$

da f. utilidade (2) temos:

$$d = x\beta + \epsilon.$$

Se assumimos que $\mathrm{E}[\epsilon \mid x] = 0$ então

$$\hat{\beta} = \left(x' V_d^{-1} x\right) x' V_d \hat{d}$$

$$\hat{\epsilon} = \hat{d} - x\hat{\beta}$$

Estimação por GLS.
$\hat{d}$ vetor de coeficientes estimados.
$V_d = \text{Var}(\hat{d})$