



Universidade Federal do Ceará

05



SÉRIE  
ESTUDOS ECONÔMICOS CAEN

# Análise de Demanda Agregada por Produtos Diferenciados

Sérgio Aquino DeSouza

FORTALEZA • MAIO • 2009

# Análise de Demanda Agregada por Produtos diferenciados

Sergio A. DeSouza\*

## Resumo

As preferências do consumidor, que formam a demanda, ocupam papel importante na decisão ótima das firmas. De fato, uma vez determinada a demanda é possível analisar algumas questões ligadas à Organização Industrial (efeitos de entrada e saída de produtos, efeitos de redução de impostos, simulação de fusões, etc.). O objetivo deste trabalho consiste em expor o conjunto de metodologias mais importantes de análise empírica da demanda por produtos diferenciados, ressaltando as vantagens e desvantagens de cada modelo. O modelo AIDS encontra vasta aplicação na estimação de modelos agregados de demanda. Possui boas propriedades (flexibilidade, o uso de técnicas econométricas tradicionais) de forma que constitui a escolha natural na maioria dos casos. No entanto, o modelo AIDS pode se tornar impraticável em situações que envolvam mercados com muitas variedades devido ao problema do excesso de parâmetros. Surge então a necessidade de utilizar os modelos de escolha discreta (Logit, Logit Agrupado e Mixed Logit). Dentre os quais os mais apropriados seriam o LA e o ML. O primeiro resolve parcialmente alguns problemas conceituais de Logit e possui a vantagem (em relação ao ML) de ser um modelo linear e, portanto, de fácil tratamento econométrico. O ML, por sua vez, se mostra um modelo que possui excelentes propriedades (gera matriz de elasticidades flexível, é capaz de lidar com mercados que possuem muitas variedades e evita a determinação a priori de agrupamento, como no caso do LA). No entanto, trata-se de modelo cuja implementação é bastante custosa em termos de tempo e esforço computacional.

\* Curso de pós-graduação em Economia (CAEN) e Departamento de Teoria Econômica, Universidade Federal do Ceará. Endereço: Av. da Universidade 2700, 2º andar, Fortaleza, CE, Brazil. Fone: 55-85-33667751; e-mail: sergiodesouza@ufc.br

## 1. Introdução

A oferta constitui o ponto central na análise antitruste, pois seu objetivo principal da consiste em compreender o comportamento das firmas (apreçamento, propaganda, entrada no mercado, etc.). No entanto, as preferências do consumidor, que formam a demanda, ocupam papel importante na decisão ótima das firmas. De fato, uma vez determinada a demanda é possível analisar algumas questões ligadas à Organização Industrial (mensurar o efeito de entrada e saída de produtos, determinação do mercado relevante, simulação de fusões e desinvestimentos). Desta forma, esforços consideráveis têm sido empreendidos por pesquisadores da área com o intuito de modelar sistemas de demanda cujos parâmetros possam ser determinados empiricamente. Este texto se restringe a analisar modelos agregados de demanda (em nível de produto) e, portanto, não inclui comentários sobre modelos desagregados de demanda (em nível do consumidor) - ver McFadden(1981).

Existem duas categorias de modelos de demanda: Modelos de Demanda por produtos homogêneos e modelos de demanda por produtos diferenciados. A primeira categoria encontra metodologias consolidadas e que, em geral, se apóiam em métodos empíricos relativamente simples. A segunda categoria, no entanto, possui uma série de desafios metodológicos que só foram (parcialmente) superados recentemente.

Existem duas classes de modelos de demanda por produtos diferenciados. A primeira é formada por modelos baseados em um consumidor representativo que atribui utilidade direta ao consumo dos bens ofertados no mercado. Neste caso, o pesquisador se depara com número de equações igual

ao número de produtos, gerando um sistema com muitos parâmetros. De fato, em cada equação, além dos deslocadores de demanda, devem ser especificados o efeito próprio (a sensibilidade da demanda do bem  $j$  em relação a seu próprio preço) e o efeito cruzado (a sensibilidade da demanda do bem  $j$  em relação ao preço do bem substituto  $r$ ). Em termos teóricos o número elevado de parâmetros não impõe severas limitações, o que não é verdadeiro para análise empírica como veremos nas seções a seguir. Exemplos notórios que se inserem nesta classe de modelos são: LES (Linear expenditure system) e AIDS (Almost Ideal Demand System) desenvolvido por Deaton and Muellbauer (1980).

A solução para este problema baseia-se na idéia originalmente proposta por Lancaster (1966), que consiste em assumir que os consumidores atribuem utilidade às características dos bens. Nesta perspectiva, o consumidor escolhe o bem que lhe confere a melhor combinação de atributos. A escolha do produto se dá de forma indireta a partir das preferências dos consumidores pelos atributos e não pelo produto em si. Os trabalhos de Lancaster encontram contrapartida em Organização Industrial (OI) empírica nos modelos de escolha discreta propostos por Berry(1994) e Berry, Levinsohn e Pakes (2005), que possuem a propriedade empiricamente desejável de reduzir o número de parâmetros. Estes modelos encontram vasta aplicação em Organização Industrial empírica.

O objetivo deste trabalho consiste em expor o conjunto de metodologias mais importantes de análise empírica da demanda por produtos diferenciados, ressaltando as vantagens e desvantagens de cada modelo.

## 2. Modelo AIDS

O modelo AIDS é bastante popular na análise empírica da demanda. Neste artigo apresenta-se apenas a versão mais simples, com forma linear e com apenas dois estágios: estágio superior e estágio inferior.

### Estágio superior

$$(1) \quad \log Q_{nt} = \alpha_{nt} + \gamma \log P_{nt} + \lambda \log X_{nt} + Z_{nt} \theta + \eta_{nt}$$

$X_{nt}$  - A renda (real) disponível dos consumidores no mercado n no período t.

$P_{nt}$  - índice de preços da indústria. Utiliza-se frequentemente a aproximação linear do índice teórico, apresentada por Stone (1954), dado por:

$$\log P_{nt} = \sum_i w_i \log p_{int}$$

Onde  $w_i$  é o share de valor médio da marca i na região n. Média é calculada ao longo dos anos disponíveis na amostra.

$Z_{nt}$  - vetor que contém outras variáveis que explicam a demanda (tendência, termos sazonais, temperatura, etc.)

$\eta_{nt}$  - Erro do estágio superior

### Estágio inferior

A share de valor de cada marca i é dada por:

$$(2) \quad s_{int} = \alpha_{in} + \beta_i \log(Y_{nt} / P_{nt}) + \sum_j \gamma_{ij} \log p_{jnt} + Z_{nt} \theta_{in} + \varepsilon_{int}$$

$Y_{nt}$  - Receita total da indústria

$s_{int}$  - Share de valor. É calculada pela razão entre receita da marca i ( $Y_{int}$ ) e receita da indústria, i.e.,  $s_{int} = Y_{int} / Y_{nt}$

$p_{jnt}$  - preço do produto j.

As elasticidades são dadas por

$$(3) \quad \eta_{ii} = \left[ \frac{1}{s_i} (\gamma_{ii} - \beta_i w_i) + \left(1 + \frac{\beta_i}{s_i}\right) (1 + \gamma) w_i \right] - 1, \text{ Elasticidade pr\u00f3pria}$$

$$(4) \quad \eta_{ij} = \frac{1}{s_i} (\gamma_{ij} - \beta_i w_j) + \left(1 + \frac{\beta_i}{s_i}\right) (1 + \gamma) w_j, \text{ Elasticidade Cruzada}$$

Estima\u00e7\u00e3o do modelo AIDS

A econometria de equa\u00e7\u00f5es simult\u00e2neas pode ser usada para estimar a regress\u00e3o acima. O m\u00e9todo mais imediato e de f\u00e1cil implementa\u00e7\u00e3o computacional \u00e9 o SURE (*Seemingly Unrelated Regression*). No entanto, a exemplo da aplica\u00e7\u00e3o de MQO para produtos homog\u00eaneos, o pre\u00e7o \u00e9 uma vari\u00e1vel end\u00f3gena devido a sua correla\u00e7\u00e3o com o erro. Neste caso o SURE gera estimadores inconsistentes. A exemplo da demanda por produto homog\u00eaneos, ignorar o problema da endogeneidade tipicamente gera subestima\u00e7\u00e3o do coeficiente da vari\u00e1vel pre\u00e7o em termos absolutos, o que resulta em subestima\u00e7\u00e3o das elasticidades e superestima\u00e7\u00e3o do poder de mercado. Esta observa\u00e7\u00e3o \u00e9 extremamente relevante para a an\u00e1lise emp\u00edrica em Organiza\u00e7\u00e3o Industrial. De fato, quanto menor a sensibilidade do consumidor em rela\u00e7\u00e3o a aumento de pre\u00e7os, isto \u00e9, quanto menor a elasticidade (em m\u00f3dulo), maior \u00e9 a capacidade de eleva\u00e7\u00e3o de pre\u00e7os por parte das firmas (maior poder de mercado). **Portanto, ignorar a endogeneidade leva a conclus\u00e3o enganosa de que as firmas possuem poder de mercado maior do que realmente det\u00eam. No contexto de defesa concorr\u00eancia, por exemplo, tal conclus\u00e3o facilitaria a ocorr\u00eancia de falsas positivas (reprova\u00e7\u00e3o quando na realidade a conduta da firma ou fus\u00f5es**

**entre firmas deveriam ser aprovados) no julgamento de casos de antitruste.**

A solução comumente adotada consiste no uso de métodos baseados em variáveis instrumentais. O método Mínimos Quadrados em Três Estágios (MQ3E) é o mais utilizado para estimar o modelo AIDS. Alternativamente poderia ser utilizado o Método Geral dos Momentos (MGM), que acomoda o MQ3E como caso particular.

Um dos principais problemas na estimação de demanda por produtos diferenciados consiste na busca por instrumentos. Deslocadores de custo, comumente utilizados em estimação de demanda por produto homogêneo raramente são utilizados no presente caso, visto que dificilmente se encontram dados sobre custos que variam entre as marcas/produtos. Uma solução é apresentada por Hausman et al. (1994) que explora a possível observação de mercados do mesmo produto, mas geograficamente distintos. O autor sugere como instrumentos preços do mesmo produto em mercados geograficamente separados. Mais especificamente, utiliza-se o preço do produto  $j$  no mercado A como instrumento para o preço do produto  $j$  no mercado B. O uso deste conjunto de instrumentos baseia-se na hipótese de que produtos com as mesmas características são fabricados com tecnologia e insumos semelhantes. Portanto, choques de custos que, por natureza, não são relacionados com choques de demanda, geram correlação dos preços do mesmo produto em mercados geograficamente distintos. Uma crítica à utilização destes instrumentos (Bresnahan, 1996) reside na possibilidade de choques agregados (que afetam todos os mercados geográficos) não observados deslocarem o preço de um mesmo produto em todos os mercados, gerando correlação dos

preços com os choques e, conseqüentemente, invalidando os instrumentos. Vale ressaltar também que o uso destes instrumentos necessita da observação de dados em mercados geograficamente distintos, informação que nem sempre está disponível para o pesquisador.

### Vantagens e Desvantagens do AIDS

Uma das principais virtudes do modelo AIDS reside na sua flexibilidade, isto é, sua capacidade de acomodar qualquer padrão de substituição entre produtos. Ou seja, o modelo AIDS não impõe a priori, em sua formulação teórica, nenhum padrão para a matriz de elasticidades. Neste sentido o modelo AIDS está mais alinhado na comparação com métodos alternativos com o princípio *Let the data talk*.

Sua principal desvantagem reside na quantidade excessiva de parâmetros a serem estimados. Por exemplo, para mercado com N produtos, no mínimo,  $(N^2+3N-4)/2$  parâmetros devem ser estimados (Hausman e Leonard, 2005). **Ou seja, no modelo AIDS existe uma relação exponencialmente crescente entre o número de produtos N e o numero de parâmetros a serem estimados.** Para mercado com poucos produtos, este problema pode não ser restritivo. No entanto, para mercados com vários produtos o modelo AIDS pode se tornar impraticável. Por exemplo, no mercado de automóveis é comum encontrar quantidade de modelos que se aproximam do número 70, o que implicaria na estimação de 2553 parâmetros.



## *Sumário de vantagens e desvantagens do Logit*

### *Vantagens*

- Uso de técnicas de regressão Linear
- Matriz de Elasticidade Flexível. Neste sentido o modelo AIDS está mais alinhado na comparação com métodos alternativos com o princípio *Let the data talk*.

### *Desvantagem*

- Excesso de parâmetros em modelos que incluem muitos produtos.

## **3. Modelos de Escolha discreta**

Conforme discussão anterior, esta classe de modelos assume que a escolha do produto se dá de forma indireta a partir das preferências dos consumidores pelos atributos e não pelo produto em si. Uma vez mapeadas as preferências do consumidor pelos atributos é possível inferir a distribuição de preferências pelos produtos. No entanto, cabe a pergunta: Qual a vantagem deste mapeamento indireto das escolhas do consumidor? A resposta é simples: o espaço de atributos usualmente possui dimensão (número de parâmetros) menor que o espaço de produtos. Logo, a principal vantagem reside na redução do número de parâmetros a serem determinados empiricamente. Os modelos de escolha discreta rompem a relação entre número de produtos e número de parâmetros, permitindo aplicação em mercados caracterizados pela presença de muitas variedades (e.g. automóveis). Não interessa o quão elevada seja a quantidade de produtos, o número de parâmetros a ser estimado será o mesmo. Isto é verdade para

todos os modelos de escolha discreta, no entanto, este resultado fica bem mais claro no caso do modelo *Logit* descrito a seguir.

### 3.1 Modelo Logit

Formalmente, neste modelo o consumidor  $i$  atribui ao produto  $j$  a seguinte utilidade  $u_{ij}$ :

$$(5) \quad u_{ij} = -\alpha p_j + x_j \beta + \xi_j + \varepsilon_{ij}$$

onde  $x_j$  representa um vetor (linha) de características dos produtos de dimensão  $K$  (no caso de veículos, tamanho, peso, potência seriam exemplos de características),  $\xi_j$  é um índice que agrupa outras características não incluídas no vetor  $x_j$  e  $\varepsilon_{ij}$  adiciona a utilidade um erro estocástico de média zero. Por conveniência expositiva, é interessante reescrever a utilidade da seguinte forma:

$$(6) \quad u_{ij} = \delta_j + \varepsilon_{ij} \quad ; \quad \text{onde } \delta_j = -\alpha p_j + x_j \beta + \xi_j$$

A especificação da demanda se completa com a definição do bem externo, ao qual o consumidor  $i$  atribui utilidade  $u_{i0} = \delta_0 + \varepsilon_{ij}$ . Assume-se que  $\delta_0 = 0$ , uma normalização típica de modelos de escolha discreta. O consumidor então escolhe o produto que lhe confere maior utilidade, o que corresponde a resolução da seguinte maximização,  $Max \{u_{ij}, j = 0, 1, \dots, N\}$ , onde  $n$  representa o número de bens internos.

Ao assumir uma distribuição de valor extremo do tipo II para  $\varepsilon_{ij}$  ( $f(\varepsilon_{ij}) = \exp(-\exp(\varepsilon_{ij}))$ ), é possível obter uma forma analítica para a probabilidade de o consumidor  $i$  escolher o determinado produto  $j$ , cuja fórmula é dada por

$$(7) \quad prob_{ij} = \frac{e^{\delta_j}}{\sum_{k=0}^I e^{\delta_k}}$$

Observe que a equação acima não é indexada pelo consumidor  $i$ . Portanto, a probabilidade não condicional, ou seja, a probabilidade do produto  $j$  ser escolhido ( $prob_j$ ) é idêntica à probabilidade condicional  $prob_{ij}$ . Uma identidade comum em modelos Logit consiste em igualar  $prob_j$  à fatia de mercado do produto  $j$  ( $s_j$ ). Logo, a demanda por  $j$  é dada por

$$(8) \quad s_j = \frac{e^{\delta_j}}{\sum_{k=0}^N e^{\delta_k}}$$

Ou, em sua forma log-linear

$$(9) \quad \ln s_j - \ln s_0 = \delta_j - \delta_0 = \delta_j = -\alpha p_j + x_j \beta + \xi_j$$

As elasticidades são derivadas são dadas por:

$$(10) \quad \frac{p_j}{s_j} \frac{\partial s_j}{\partial p_j} = -\alpha p_j [1 - s_j] \quad ; \text{ Elasticidade própria}$$

$$(11) \quad \frac{p_r}{s_j} \frac{\partial s_j}{\partial p_r} = -\alpha p_r s_r \quad ; \quad \text{Elasticidade cruzada, (j diferente de r)}$$

### *Estimação do Modelo Logit*

As informações necessárias para implementar o modelo Logit são :

- Quantidades produzidas (ou vendidas)  $q_j$ 's
- Tamanho do Mercado  $M$  (número de potenciais consumidores do produto definido de forma agregada). Por exemplo, no caso de mercado de automóveis novos, seria o número de potenciais compradores do produto automóvel 0Km.
- Preços dos produtos  $p_j$ 's
- Características dos produtos  $x_j$  's

De posse dos  $q_j$  's e  $M$  é possível calcular as fatias de mercado  $s_j = q_j / M$  .

Uma distinção importante entre os modelos de escolha discreta e o modelo AIDS reside na definição de share(fatia ou participação). No modelo AIDS, a share de um produto  $j$  corresponde a participação do valor monetário das vendas deste produto em relação ao valor monetário agregado das vendas totais da indústria. Nos modelos de escolha discreta, a share de um produto  $j$  corresponde a participação da quantidade vendida em relação quantidade tamanho do mercado ( $M$ ). Utilizam-se os termos share de valor (AIDS) e share de quantidade (modelos de escolha discreta). O próximo passo consiste em estimar a seguinte equação:

$$(12) \quad \ln s_j - \ln s_0 = -\alpha p_j + x_j \beta + \xi_j$$

A econometria tradicional pode ser usada para estimar a regressão acima. O método mais imediato e de fácil implementação computacional é o MQO. No entanto, a exemplo da equação para produtos homogêneos e para o modelo AIDS, preço é uma variável endógena devido a sua correlação com o erro, índice que sumariza os elementos que deslocam a demanda relativa do bem  $j$ , mas que não estão incluídos no vetor  $x_j$ . Desta forma, ignorar o problema da endogeneidade tipicamente gera subestimação do coeficiente da variável preço em termos absolutos, o que resulta em subestimação das elasticidades e superestimação do poder de mercado.

Uma solução seria usar algum método que utiliza variáveis instrumentais (VI). No caso de modelos de escolha discreta, existem duas classes de variáveis instrumentais: VI's construídas a partir da observação de preços do mesmo produto em mercados geograficamente distintos (Hausman et. al, 1994) conforme descrição anterior e VI's construídas a partir das características dos produtos, propostas por Berry, Levinsohn e Pakes (1995), doravante denominado BLP.

As VI's de BLP são:

- As próprias características do produto;
- As somas das características dos outros produtos produzidos pela mesma firma;
- As somas das características dos produtos produzidos pelas outras firmas.

As VI's de Hausman são:

- As próprias características do produto;
- Preço do produto  $j$  no mercado A como instrumento para o preço do produto  $j$  no mercado B

Em todos os modelos de escolha discreta, assume-se que as características são exógenas, daí a inclusão no rol dos instrumentos. Ou seja, o regressor é instrumento dele próprio. A exogeneidade das características é uma hipótese discutível, visto que é de se esperar que muitos atributos sejam correlacionados. Por exemplo, veículos de luxo usualmente possuem elevada potência (característica usualmente incluída nos regressores) e design mais arrojado ou melhor acabamento (características usualmente excluídas do conjunto de regressores pois raramente são observadas pelo pesquisador).

### *Problemas Conceituais do Modelo Logit*

Mesmo de posse de bons instrumentos e de estimadores com propriedades econométricas desejáveis (coeficientes significantes e com sinal correto) o modelo Logit pode não ser adequado para análise econômica (cálculo das elasticidades, simulação de fusões, etc.) em muitos casos, pois apresenta sérios problemas conceituais. As limitações conceituais do Logit se revelam claramente em duas dimensões: cálculo da margem preço-custo e do grau de substituição entre produtos.

De acordo com o modelo Logit a margem preço custo, definida neste trabalho como a diferença entre estas duas variáveis, é dada por:

$$(13) \quad p_j - c_j = 1/(\alpha[1 - s_j])$$

Note que, independente da estimação de  $\alpha$ , dois produtos que possuem a mesma share induzem a mesma margem, o que nos leva ao paradoxo CIVIC-KA. No ano de 2008 as vendas de Honda CIVIC e FORD KA foram aproximadamente as mesmas no mercado brasileiro (68000 unidades de cada modelo), o que implica em fatias de mercado iguais. Portanto, o modelo Logit implica que a margem do FORD KA é igual à margem do HONDA CIVIC. A restrição do modelo Logit fica mais clara na medida da margem definida anteriormente. No entanto, conclusões semelhantes podem ser estendidas ao índice de Lerner (margem/preço), medida da capacidade de elevar preços mais comumente utilizada (Nevo, 2000a e Huse e Salvo, 2006).

Com relação ao grau de substituição entre dois produtos  $j$  e  $r$ , observe que, a partir das elasticidades cruzadas  $\frac{p_r}{s_j} \frac{\partial s_j}{\partial p_r} = -\alpha_{p_r, s_r}$ , o aumento da fatia de mercado do produto  $j$  decorrente do aumento percentual do preço  $p_r$  depende apenas de  $r$ . Isto significa que um aumento percentual de  $p_r$  afetará de forma idêntica todos os outros produtos no mercado (competição não localizada). Trata-se de uma propriedade pouco plausível em mercados com produtos diferenciados. Com efeito, uma das motivações principais para estudar este tipo de mercado é justamente, para cada par de produtos, distinguir entre mais próximos e mais distantes em relação ao grau de diferenciação. O paradoxo CIVIC-KA aparece novamente. De fato, de acordo com o modelo Logit, em termos percentuais, um aumento no preço do FORD KA afetará de forma idêntica a demanda do CELTA e a demanda do HONDA CIVIC. Estas limitações do Logit são manifestações da propriedade de independência de alternativas irrelevantes, amplamente discutida em modelos de demanda desagregados (em nível do consumidor). A solução mais simples para esta restrição do modelo Logit é oferecida pelo Modelo Logit Agrupado (ou *Nested Logit Model*). Neste caso, o pesquisador define *a priori* os agrupamentos (ou segmentos do mercado) e assume que produtos pertencentes ao mesmo grupo possuem grau de substituição mais elevado. Outra solução, bem mais sofisticada se refere ao modelo Logit com coeficientes aleatórios (Random Coefficients models), onde a utilidade marginal pelos atributos, incluindo preços, varia de consumidor para consumidor.



## *Sumário de vantagens e desvantagens do Logit*

### *Vantagens*

- Uso de técnicas de regressão Linear
- Número reduzido de parâmetros. Ou seja, Logit pode lidar com mercados caracterizados pela presença muitas variedades.
- Um vez observadas, as características podem servir como instrumentos (BLP).

### *Desvantagens*

- Padrão de substituição pouco razoável e definido a priori.
- Resultados pouco plausíveis com relação à medição da margem preço-custo.
- Baixíssimo grau de Flexibilidade.
- Matriz de Elasticidades com vários elementos idênticos.
- Neste sentido o modelo LOGIT está distante do princípio *Let the data talk*. Está mais para *Let the Model Do Everything*.

### *3.2 Logit Agrupado*

O modelo Logit Agrupado (LA) supõe que a decisão de compra dos consumidores é racionalizada por meio de uma árvore de decisão onde cada nível representa escolhas tomadas dentro de grupos e subgrupos de alternativas.

A propriedade de independência de alternativas irrelevantes (IAI) presente em modelos *logit* é corrigida parcialmente. Ela permanece válida para as alternativas existentes dentro dos subgrupos mais restritos, mas geralmente não será válida para bens pertencentes a diferentes grupos.

Este modelo é adequadamente aplicado a vários mercados onde é razoável supor que os consumidores percebem certas classes (grupos ou ninhos) de produtos, onde em cada classe estão contidos variedades com alto grau de substituição entre si. Estas classes podem ser diretamente representadas por segmentações existentes no mercado. O mercado de automóveis pode ser segmentado, por exemplo, de acordo com a classe do automóvel (compacto, médio, esportivo, van, luxo, etc), origem (nacional ou importado).

Nesta forma de modelagem seguimos a mesma estrutura de utilidade indireta condicional obtida pelo consumidor  $i$  na compra de um carro  $j$ ,  $u_{ij} = \delta_j + v_{ij}$ , onde o nível médio de utilidade do produto  $j$  é dado por  $\delta_j = -\alpha p_j + \beta X_j + \xi_j$ . No entanto, temos agora que o termo  $v_{ij}$  pode ser decomposto do seguinte modo:

$$(14) \quad v_{ij} = \zeta_{ig} + (1 - \sigma)\epsilon_{ij}$$

Assim temos a seguinte função utilidade indireta condicional do indivíduo

$i$ :

$$(15) \quad u_{ij} = \delta_j + \zeta_{ig} + (1 - \sigma)\epsilon_{ij}$$

O primeiro termo aleatório  $\zeta_{ig}$  representa o efeito de “choques” que afetam todos os produtos de um determinado grupo  $g$  e sua distribuição depende do parâmetro  $\sigma$  ( $0 \leq \sigma < 1$ )<sup>1</sup>. O segundo termo aleatório  $\varepsilon_{ij}$  é assumido identicamente e independentemente distribuído de acordo com uma distribuição de valor extremo tipo II vista anteriormente.

De acordo com Berry (1994), podemos interpretar a utilidade descrita acima como um modelo de coeficientes aleatórios  $\zeta_{ig}$ , onde estes coeficientes operam apenas em variáveis *dummy* específicas.

Definindo-se  $d_{jg}$  como uma variável *dummy* tomando o valor 1 quando  $j$  pertence ao grupo  $g$ , podemos escrever a equação anterior da seguinte forma:

$$(16) \quad u_{ij} = \delta_j + \sum_g [d_{jg} \cdot \zeta_{ig}] + (1 - \sigma)\varepsilon_{ij}$$

Ainda de acordo com Berry (1994), assumindo uma forma funcional para a função utilidade e assumindo que  $v_{ij}$  tem uma distribuição de valor extremo<sup>2</sup> podemos obter uma forma analítica para a função de *market share*. De fato, é possível demonstrar que o *market share* do produto  $j$  de um determinado grupo  $g$  é dada por:

$$(17) \quad s_{j|g} = \frac{e^{\delta_j/(1-\sigma)}}{D_g}$$

, onde  $D_g = \sum_{j \in J_g} e^{\delta_j/(1-\sigma)}$ .

---

<sup>1</sup> Conforme o valor de  $\sigma$  aproxima-se de 1, a correlação dos níveis de utilidade dentro do grupo tende a 1 e, conforme seu valor se aproxima de 0, tal correlação tende a 0.

<sup>2</sup> Cardel mostra que se  $\mathcal{E}$  possui uma distribuição de valor extremo, então o termo  $[\zeta + (1 - \sigma)\varepsilon]$  também possui esta forma de distribuição [Cardell (1991) *apud* Berry (1994)].

O market share do grupo g é dado por

$$(18) \quad s_g = \frac{D_g^{(1-\sigma)}}{\sum_g D_g^{(1-\sigma)}}$$

E a fatia de mercado do carro j no mercado é dada por

$$(19) \quad s_j = s_{j/g} s_g = \frac{e^{\delta_j/(1-\sigma)}}{D_g^\sigma \sum_g D_g^{1-\sigma}}$$

Seguindo Berry (1994), a equação acima pode ser invertida para ser transformada em uma equação de demanda linear :

$$(20) \quad \ln s_j - \ln s_0 = -\alpha p_j + X_j \beta + \sigma \ln s_{j/g} + \xi_j$$

Onde  $s_0$  é a proporção de consumidores que escolhem a alternativa externa (ou seja, não comprar o produto interno) e  $\alpha$ ,  $\beta$ ,  $\sigma$  são parâmetros a serem estimados.

Como o modelo é linear, os métodos empregados para o Logit são mesmos do LA com os mesmos instrumentos. O modelo LA apresenta avanços em relação o Logit. De fato, após manipulações algébricas, obtêm-se as seguintes fórmulas:

$$(21) \quad \frac{p_j \partial s_j}{s_j \partial p_j} = -\frac{\alpha}{1-\sigma} p_j \{1 - [(1-\sigma)s_j + \sigma s_{j/g}]\} ; \text{ Elasticidade própria}$$

$$(22) \quad \frac{p_r \partial s_j}{s_j \partial p_r} = \frac{\alpha}{1-\sigma} p_r [(1-\sigma)s_r + \sigma s_{r/g}]; \text{ ELASTICIDADE CRUZADA, j e r}$$

pertencem ao mesmo grupo

$$(23) \quad \frac{p_r}{s_j} \frac{\partial s_j}{\partial p_r} = \alpha p_r s_r; \text{ ELASTICIDADE CRUZADA, } j \text{ e } r \text{ pertencem a grupos}$$

diferentes

O Logit Agrupado apresenta avanços em relação ao Logit, que se revelam claramente a partir das elasticidades que o modelo implica. Mostra-se que no modelo Logit Agrupado:

$$(24) \quad p_j - c_j = \frac{1}{\frac{\alpha}{1-\sigma} \{1 - [(1-\sigma)s_j + \sigma s_{j/g}]\}}$$

Note que dois produtos com a mesma share não teriam necessariamente a mesma margem (p-c), pois a margem não depende apenas da share do produto no mercado como um todo. Depende também da share do produto no agrupamento a que pertence. Na realidade depende da média ponderada, tendo  $\sigma$  como fator de ponderação  $[(1-\sigma)s_j + \sigma s_{j/g}]$ . Quanto maior  $\sigma$ , maior será o peso da presença do produto em seu segmento ( $s_{j/g}$ ) e menor será a peso de sua presença no mercado como um todo ( $s_j$ ). Trata-se de um resultado com interpretação bastante intuitiva. Note que a margem (p-c) é diretamente proporcional ao termo ponderado  $[(1-\sigma)s_j + \sigma s_{j/g}]$ . Portanto, é possível concluir um resultado com interpretação bastante intuitiva: quanto maior for a importância da segmentação (valor elevado de  $\sigma$ ) maior será importância da presença do produto no segmento ( $s_{j/g}$ ) em relação à sua presença no mercado como um todo ( $s_j$ ) para explicar a margem auferida. No caso Ka-Civic, ambos apresentam a mesma share ( $s_j$ ) pois vendem nas

mesma quantidade, mas não necessariamente a mesma share em seu respectivo segmento  $s_{j/g}$ . Neste caso é de se esperar que o Civic tenha uma presença maior em seu segmento em relação a presença que o Ford Ka possui em seu segmento. Afinal 68000 unidades no mercado de sedans (possível segmentação) devem significar bem mais que a mês quantidade no mercado de populares. Assim, é possível demonstrar que o modelo Logit Agrupado implica em uma margem do FORD KA menor que a margem do HONDA CIVIC.

Com relação à substituição entre produtos o LA também apresenta avanços em relação ao Logit. Observe, a partir das elasticidades cruzadas (equações (22) e(23) que o aumento da fatia de mercado do produto  $j$  decorrente do aumento percentual do preço  $p_r$  **não** depende apenas de  $r$ . Depende também do grupo  $a$  que pertence. Isto significa que um aumento percentual de  $p_r$  **não** afetará de forma idêntica todos os outros produtos no mercado. De fato, é possível demonstrar que um aumento no preço de  $p_r$  afetará mais intensamente a demanda por um bem pertencente ao mesmo grupo em relação a demanda por um bem em outro grupo.

Observe que o modelo Logit Agrupado gera um padrão de substituição entre os produtos mais plausível que aquele gerado pelo Logit. O grau de substituição entre dois produtos depende do fato de pertencerem ou não ao mesmo segmento. Portanto, o LA é mais flexível que o modelo Logit. No entanto, note na comparação com o modelo AIDS, o modelo LA ainda apresenta baixo grau de flexibilidade. De fato, note que, dentro de cada segmento, as elasticidades cruzadas serão idênticas. Ou seja, matriz de elasticidades ainda apresenta alguns elementos idênticos.

## Sumário das vantagens e desvantagens do LA

### Vantagens:

- Uso de técnicas de regressão Linear
- Número reduzido de parâmetros. Ou seja, Logit Agrupado pode lidar com mercados caracterizados pela presença muitas variedades.
- Uma vez observadas, as características podem servir como instrumentos (BLP).
- Padrão de substituição entre os produtos mais razoável que aquele gerado pelo Logit.
- Em relação ao Logit, gera medidas mais plausíveis de margem (poder de mercado)

### Desvantagens:

- Grau de flexibilidade ainda baixo em relação ao AIDS.
- Matriz de elasticidades com alguns elementos idênticos.
- Neste sentido o modelo LA ainda está distante do princípio *Let the data talk*. Está mais para *Let the Model Do Almost Everything*.

### 3.3 Logit com Coeficientes Aleatórios ou Mixed Logit (Artigo BLP)

Formalmente, nestes modelos o consumidor  $i$  atribui ao produto  $j$  a seguinte utilidade  $u_{ij}$ :

$$(25) \quad u_{ij} = -\alpha_i p_j + \sum_k x_{jk} \beta_{ik} + \xi_j + \varepsilon_{ij}$$

onde  $x_j$  representa um vetor (linha) de características dos produtos de dimensão  $K$  (no caso de veículos, tamanho, peso, potência seriam exemplos

de características),  $\xi_j$  é um índice que agrupa outras características não incluídas no vetor  $x_j$ <sup>3</sup> e  $\varepsilon_{ij}$  adiciona a utilidade um erro estocástico de média zero. Por conveniência expositiva, é interessante reescrever a utilidade da seguinte forma:

Os coeficientes  $\alpha_i$  e  $\beta_{ik}$  são aleatórios e seguem uma determinada distribuição. A exposição a seguir se aproxima daquela encontrada em Berry, Levinsohn and Pakes (1999). Onde  $\alpha_i = \alpha / v_{yi}$ , onde o denominador corresponde à renda de cada consumidor cuja distribuição é log-normal por hipótese, e assume-se também que  $\beta_{ik}$  segue uma normal padrão com média  $\bar{\beta}_k$  e desvio padrão  $\sigma_k$ . Desta forma a utilidade pode se reescrita da seguinte forma:

$$(26) u_{ij} = -\frac{\alpha}{v_{yi}} p_j + \sum_k \bar{\beta}_k x_{jk} + \sum_k \sigma_k v_{ik} x_{jk} + \xi_j + \varepsilon_{ij}$$

Ou, ao definir  $v_i = (v_{yi}, v_{i1}, v_{i2}, \dots, v_{ik}, \dots)$ ,  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_k, \dots)$  e  $x_j = (x_{j1}, x_{j2}, \dots, x_{jk}, \dots)$

$$(27) u_{ij} = \mu_{ij}(p_j, x_j, v_i, \alpha, \sigma) + \delta_j + \varepsilon_{ij}$$

$$\text{Onde } \mu_{ij}(p_j, x_j, v_i, \alpha, \sigma) = -\frac{\alpha}{v_{yi}} p_j + \sum_k \sigma_k v_{ik} x_{jk} \text{ e } \delta_j = \sum_k \bar{\beta}_k x_{jk} + \xi_j$$

Ao assumir uma distribuição de valor extremo do tipo II para  $\varepsilon_{ij}$  ( $f(\varepsilon_{ij}) = \exp(-\exp(\varepsilon_{ij}))$ ), é possível obter uma forma analítica para a

---

<sup>3</sup> O objetivo da divisão do vetor de características em  $x_j$  e  $\xi_j$  se tornará claro na seção empírica.



probabilidade de o consumidor  $i$  escolher o determinado produto  $j$ , cuja fórmula é dada por

$$(28) \quad prob_{ij} = \frac{e^{\mu_{ij}(p_j, x_j, v_i, \alpha, \sigma) + \delta_j}}{\sum_{j=0}^1 e^{\mu_{ij}(p_j, x_j, v_i, \alpha, \sigma) + \delta_k}}$$

Portanto a probabilidade não condicional da escolha do produto  $j$  é

$$(29) \quad prob_j = s_j(p, x, \alpha, \sigma, \delta) = \int \frac{e^{\mu_{ij}(p_j, x_j, v_i, \alpha, \sigma) + \delta_j}}{\sum_{j=0}^N e^{\mu_{ij}(p_j, x_j, v_i, \alpha, \sigma) + \delta_j}} dF(v_i)$$

A equação acima representa um sistema de equações de demanda por produtos diferenciados, pois depende de preços e características de todos os produtos do mercado. No entanto o alto grau de não linearidade nos regressores, e principalmente no erro  $\xi_j$ , impede o uso imediato de técnicas econométricas tradicionais (lineares ou não<sup>4</sup>). A estratégia sugerida por BLP consiste em isolar o erro em função dos parâmetros do modelo e construir momentos empíricos a partir do uso de variáveis instrumentais, utilizando MGM. Abaixo segue versão simplificada do algoritmo econométrico sugerido por BLP

---

<sup>4</sup> Métodos não-lineares tradicionais também não podem ser utilizados diretamente, visto que tais modelos acomodam apenas a possibilidade de os regressores entrarem de forma não linear na equação. O erro continua a compor a regressão na forma linear neste modelos, o que não é o caso do sistema de demanda exposto na equação (29).

## Algoritmo

- Passo1 (Loop externo) : Iniciar com valores dos parâmetros  $(\alpha^0, \sigma^0, \bar{\beta}^0)$
- Passo 2 (Loop interno): Resolver o sistema de equações abaixo em relação a  $\delta_j$ 's

$$\bar{s}_j = \int \frac{e^{\mu_{ij}(p_j, x_j, v_i, \alpha, \sigma) + \delta_j}}{\sum_{j=0}^N e^{\mu_{ij}(p_j, x_j, v_i, \alpha, \sigma) + \delta_k}} dF(v_i), \text{ onde } \bar{s}_j \text{ representa a fatia de mercado}$$

observada produto j. A resolução deste sistema tem como resultado  $\hat{\delta}_j$ 's

- Passo 3: Obter os erros  $\xi_j = \hat{\delta}_j - \sum_k \bar{\beta}_k x_{jk}$
- Passo 4: Construir função objetivo MGMA partir dos momentos empíricos

$$Func\_Objetivo = (\sum_j \xi_j z_{j1})^2 + (\sum_j \xi_j z_{j2})^2 + (\sum_j \xi_j z_{j3})^2 + \dots (\sum_j \xi_j z_{jL})^2 + \dots (\sum_j \xi_j z_{jL})^2$$

Onde L denota o número de instrumento e  $z_{jl}$  representa um dos

instrumentos utilizados

### *Prós e contras do Mixed Logit*

As vantagens no uso do Mixed Logit são :

- Número reduzido de parâmetros. Ou seja, a exemplo dos outros modelos de escolha discreta o Mixed Logit pode lidar com mercados caracterizados pela presença muitas variedades.
- Uma vez observadas, as características podem servir como instrumentos (BLP).

-Modelo não impõe a priori um padrão de substituição entre os produtos.

-Em relação ao Logit e ao LA, Mixed Logit gera medidas mais plausíveis de poder de mercado.

-Modelo é bastante Flexível. Matriz de elasticidades não possui forçosamente valores idênticos.

Desvantagens do uso do Mixed Logit são:

-Impossibilidade do uso de técnicas usuais de regressão (Linear ou não lineares). Modelo não está disponível em pacotes econométricos.

-Modelo possui carga computacional relativamente elevada.

#### 4. Comentários finais

Segue abaixo quadro comparativo que sumariza propriedades das diferentes metodologias expostas neste artigo

**Quadro 1: Avaliação de Metodologias de Estimação de Demanda por produtos Diferenciados**

	<b>AIDS</b> Deaton e Mull.,1980	<b>LOGIT</b> Berry,1994	<b>LOGIT AGR.</b> Berry, 1994	<b>MIXED LOGIT</b> BLP,1995
<b>Flexibilidade</b>	Alta	Muito baixa	Baixa	Alta
<b>Capacidade para lidar com muitos produtos</b>	Baixa	Alta	Alta	Alta
<b>Velocidade Analítica</b>	Alta	Alta	Alta	Baixa

O modelo AIDS encontra vasta aplicação na estimação de modelos agregados de demanda. Possui boas propriedades (flexibilidade da matriz de elasticidades, permite o uso de técnicas econométricas usuais e, conseqüentemente, a fácil de implementação através de pacotes econométricos de grande circulação) e, desta forma, constitui a escolha natural na maioria dos casos. No entanto, o modelo AIDS pode se tornar impraticável em situações que envolvam mercados com muitas variedades devido ao

problema de identificação de número excessivo de parâmetros. Surge então a possibilidade de utilizar os modelos de escolha discreta (Logit, Logit Agrupado e Mixed Logit). Dentre os quais os mais apropriados seriam o LA e o ML. O primeiro resolve parcialmente alguns problemas conceituais de Logit e possui a vantagem (em relação ao ML) de ser um modelo linear e, portanto, de fácil tratamento econométrico, através do uso de econometria tradicional. O ML, por sua vez, se mostra um modelo que possui excelentes propriedades (gera matriz de elasticidades flexível, é capaz de estimar demanda com muitas variedades e evita a determinação a priori de agrupamento, como no caso do LA). No entanto, trata-se de modelo cuja implementação é bastante custosa em termos de tempo e esforço.

## *Bibliografia*

- Berry, S. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal*, 25(2), pp. 242-262.
- Berry, S., Levinsohn, J., and Pakes, A. (1995), "Automobile Prices in Market Equilibrium" *Econometrica*, 63(4), pp. 841-890.
- Bresnahan, T. (1996). Comments on "Valuation of New Goods Under Perfect and Imperfect Competition" by Jerry Hausman, in T. Bresnahan and R. Gordon, eds., *The Economics of New Goods, Studies in Income and Wealth Vol. 58*, Chicago: National Bureau of Economic Research.
- Deaton, A. and J. Muellbauer (1980) "An Almost Ideal Demand System", *American Economic Review*, 70, 312-326.
- DeSouza, S. (2009) "Antitrust Mixed Logit Model", Mimeo, CAEN,UFC.
- Hausman, J. e Leonard, G.(2005) "Competitive analysis using a flexible demand specification" *Journal of Competition Law and Economics*" 1(2), 279-301
- Huse, C. e A. Salvo (2006). "Estimação e identificação de Demanda e de Oferta." In *Métodos Quantitativos em Defesa da Concorrência e Regulação Econômica*, Eds. Eduardo P. S. Fiuza and Ronaldo Seroa da Motta, Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada.
- Lancaster, K., 1966, "A New Approach to Consumer Theory," *Journal of Political Economy*, 74, 132–157.
- McFadden, D. (1981) "Econometric Models of Probabilistic Choice", in C. Manski and D. McFadden (Eds), *Structural Analysis of Discrete Data*.
- Hausman, J., Leonard, G., Zona J. (1994) "Competitive Analysis with Differentiated Products" *Annales d'Economie et de Statistique*, 34, 159-180
- Nevo A.(2000a) "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand". *Journal of Economics & Management Strategy*, 9(4), pp.513–548
- \_\_\_\_\_ (2000b). "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *Rand Journal of Economics*, 31, 395-421
- \_\_\_\_\_ (2001) "Measuring Market Power in the Ready-to-Eat Cereal Industry". *Econometrica*, 69(2) pp.307-342.

- Petrin, A. (2002) "Quantifying the benefits of New Products: The Case of the Minivan". *Journal of Political Economy*. 110 (4), pp. 705-729.
- Salvo, A. (2006) "Testing for Heterogeneous Business Practices across Firms in Developing Countries: The Case of the Brazilian Soft Drink Industry", Working Paper #80, Center for the Study of Industrial Organization, Northwestern University.
- Stone, J. (1954) "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand", *Economic Journal*, 64, 511-527.