

Documento de Trabalho No. 001/09

Análise de Demanda

Versão Pública

DEE

GTME

(GT n.º 3)

Este é um trabalho conjunto do Departamento de Estudos Econômicos (DEE) e do Grupo Técnico em Métodos em Economia (GTME) – GT 3.

O texto foi elaborado em conjunto por (em ordem alfabética)

Beatriz Leal (Gab.PFA/GTME)

Eduardo Pontual Ribeiro (DEE)

Gerson Carvalho (Gab. OZC/GTME)

Helenilka Pereira (Gab. FMF/GTME)

Ricardo Medeiros (Gab. VMC/GTME)

Sergio Aquino de Souza (DEE)

Washington Baldez (DEE)

O trabalho contou também com a participação dos ex-funcionários do CADEe Alexandre Lauri Henriksen (Gab. PFA/GTME) e Daniel Carvalho (Gab.PFA/GTME)

Sumário

1. Introdução	4
2. Modelos de Demanda por Produtos Homogêneos	6
3. Modelos de Demanda por Produtos Diferenciados	9
3.1. Modelo AIDS	10
3.1.1. Estimação do modelo AIDS	12
3.1.2. Vantagens e Desvantagens do Modelo AIDS	13
3.2. Modelos de Escolha Discreta	13
3.2.1. Modelo Logit	14
3.2.1.1. Estimação do Modelo Logit	16
3.2.1.2. Problemas Conceituais do Modelo Logit	17
3.2.1.3. Sumário de Vantagens e Desvantagens do Logit	18
3.2.2. Logit Agrupado	19
3.2.2.1. Vantagens e Desvantagens do Logit Agrupado	22
3.2.3. Logit com Coeficientes Aleatórios ou Mixed Logit (Artigo BLP)	22
3.2.3.1. Prós e contras do modelo Mixed Logit (ML)	24
3.3. Recomendações sobre o uso de Metodologias	24
4. Síntese Crítica dos Estudos de Casos	26
5. Comentários finais e Recomendações	29
6. Bibliografia	33

1. Introdução

Em qualquer curso de princípios de Economia, aprende-se a entender o comportamento de preços de mercado e quantidades vendidas pela interação entre oferta e demanda.

Na análise antitruste o lado da oferta constitui um ponto central, pois sintetiza as variáveis que condicionam o comportamento das firmas (apreçamento, propaganda, entrada no mercado, estrutura de custos etc.) no relacionamento com consumidores. Por outro lado, ao tomar suas decisões ótimas as firmas devem considerar as preferências e a renda do consumidor, variáveis que geram a demanda. Por exemplo, consumidores mais sensíveis a preço (demanda) não permitem às firmas cobrarem preços muito acima de seus custos (marginais).

Além do interesse empresarial de conhecer as características principais da demanda da indústria e de cada empresa, a determinação da demanda, e sua estatística síntese, a elasticidade, servem de insumo para várias etapas na análise de defesa da concorrência, como por exemplo, a determinação do mercado relevante, o grau de rivalidade e a simulação dos impactos das fusões.

Na determinação do mercado relevante, sob a estrutura conceitual do chamado “Teste do Monopolista Hipotético”, explicado no Guia para Análise Econômica de Atos de Concentração (SEAE/SDE, 2001), obter os valores das elasticidades de demanda nos permite avaliar a sensibilidade da demanda frente a aumentos de preços. Para a simulação do impacto das fusões, as elasticidades da demanda, assim como *market shares* e custos (marginais) são insumos básicos para a elaboração de tais estudos.¹

Dessa forma, esforços consideráveis têm sido empreendidos por pesquisadores da área com o intuito de modelar sistemas de demanda, cujos parâmetros possam ser determinados empiricamente. A referência clássica da literatura é o artigo de Berry

¹ A delimitação de mercado relevante e simulações de atos de concentração serão tema de textos para discussão futuros. Vale notar que para a simulação de fusões, os insumos, na realidade, são os parâmetros que definem a elasticidade, visto que esta grandeza em geral se altera com os preços.

(1994). Recentemente, temos as sínteses de DeSouza (2009), Huse e Salvo (2006) e Hosken et al. (2002) apresentando diferentes métodos que podem ser empregados na estimação da demanda². Este trabalho complementa os anteriores pois, além de expor o conjunto de metodologias mais importantes da análise empírica da demanda, destaca o uso destes métodos como ferramenta fundamental do Sistema Brasileiro de Defesa da Concorrência (SBDC). Resenhamos o uso destes métodos em atos de concentração selecionados, seja por parte dos órgãos públicos do SBDC, seja pelas partes dos processos, ilustrando a aplicação concreta dos procedimentos.

Existem duas categorias de modelos de demanda: modelos de demanda por produtos homogêneos e modelos de demanda por produtos diferenciados³. A primeira categoria encontra metodologias consolidadas e que, em geral, se apoiam em métodos empíricos relativamente simples. A segunda categoria, no entanto, possui uma série de desafios metodológicos que só foram (parcialmente) superados recentemente. Ambas as categorias serão tratadas neste documento.

Ao final da análise, destacamos recomendações para o uso prático dos métodos descritos, em particular, suas recomendações de uso em diferentes situações, envolvendo o tipo do produto em análise e disponibilidade de dados. Espera-se com isso auxiliar o SBDC e os pareceristas que atuam junto ao sistema na elaboração de estudos de demanda, peças importantes para a análise de atos de concentração e processos administrativos.

O trabalho está organizado em quatro seções além desta introdução. Inicialmente tratamos da demanda por produtos homogêneos. A demanda por produtos homogêneos constitui uma boa oportunidade para discutir os problemas básicos de estimação, que consiste na endogenia de preços e quantidades observadas. Posteriormente, trata-se da demanda por produtos heterogêneos. A quarta seção traz os estudos de casos de processos do SBDC com estimativas de demanda e a quinta seção conclui.

² A apresentação das metodologias constantes deste Documento de Trabalho parafraseia em grande parte DeSouza (2009).

³ 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), exceto no caso de modelos do tipo *cojoint analysis* - ver McFadden (1981). Em adição modelos dinâmicos, não abordados neste trabalho, são também aplicados com frequência.

2. Modelos de Demanda por Produtos Homogêneos

Um mercado caracterizado por produtos homogêneos é aquele onde o consumidor percebe todos os bens como idênticos em todas as dimensões (características intrínsecas e dimensão geográfica). Exemplos são: petróleo, gás, produtos agrícolas, gasolina. Uma equação de demanda por produto homogêneo possui as seguintes propriedades/características:⁴

- o preço é único: em cada mercado existe apenas um preço;⁵
- existe apenas uma equação para cada mercado;
- deve incluir quantidade consumida, preço, deslocadores da demanda (renda, por exemplo) e parâmetros que relacionam estas variáveis;
- deve incluir termo de erro que sumaria tudo que determina a demanda, mas que não é observado pelo pesquisador (pode também representar erros de medida).

A equação de demanda em determinado mercado n assume a seguinte forma geral, a qual acomoda modelos bastante utilizados, como o linear e o log-linear:

$$G(Q_n, P_n, \varepsilon_n, Z_n, \Omega) = 0$$

Onde:

P_n – preço

Q_n – quantidade

Z_n - vetor de deslocadores da demanda

ε_n - termo de erro que sintetiza fatores não incluídos no modelo

Ω - vetor de parâmetros do modelo

⁴ A relação entre preferências de um consumidor e sua demanda individual, assim como as propriedades matemáticas de uma função de demanda individual, e a possibilidade de agregação pode ser revista em Varian (1992) ou Jehle and Reny(2000).

⁵ Em tese, existe apenas um preço para o produto homogêneo. No entanto, na prática, alguma dispersão de preços é verificada. Se a dispersão for muito elevada, o pesquisador deve rever a hipótese de produto homogêneo. Caso contrário, o preço médio pode ser tomado como representativo e servir de input para a análise empírica.

De fato, o modelo linear pode ser escrito da seguinte forma:

$$Q_n = \alpha_0 - \alpha_Q P_n + \alpha_Z Z_n + \varepsilon_n$$

O resultado mais importante da análise de demanda consiste na estimação da elasticidade, que no caso da demanda linear é dada por

$$\varepsilon = \frac{\Delta Q}{\Delta P} \frac{P}{Q} = \alpha_Q \cdot \frac{P}{Q}$$

Como a elasticidade nesse modelo varia ao longo da curva de demanda, em geral calcula-se a elasticidade média, baseado em preços e quantidades médias entre um conjunto de pontos de quantidade e preço relevantes para a análise.

$$\varepsilon_{med} = \frac{\Delta Q}{\Delta P} \frac{P_{med}}{Q_{med}} = \alpha_Q \cdot \frac{P_{med}}{Q_{med}}$$

O modelo log-linear, por sua vez, pode ser escrito da seguinte forma:

$$\log Q_n = \alpha_0 - \alpha_Q \log P_n + \alpha_Z Z_n + \varepsilon_n$$

A elasticidade derivada do modelo log-linear é constante⁶ e igual a

$$\varepsilon = \frac{\Delta Q}{\Delta P} \frac{P}{Q} = \frac{\Delta \ln Q}{\Delta \ln P} = \alpha_Q$$

Dada a linearidade nos parâmetros das curvas de demanda comumente utilizadas para modelar mercados com produtos homogêneos, pode-se utilizar a econometria tradicional de fácil execução em softwares comumente disponíveis (STATA, EVIEWS ou Gretl).

⁶ O pesquisador deve justificar o uso de modelos de elasticidade constante, visto que trata-se de uma hipótese pouco plausível

Como em quase toda regressão linear o pesquisador pode se sentir tentado a aplicar o popular método dos Mínimos Quadrados Ordinários (MQO). No entanto, a estimação de demanda não é consistente com uma das hipóteses nas quais se baseia o método MQO, a saber, a ausência de correlação entre o erro, aquilo que o pesquisador não inclui no modelo, e as outras variáveis do lado direito da equação.

Efetivamente, choques não modelados ou observados do lado da demanda afetam simultaneamente preços e quantidades. Por exemplo, suponha a estimação de um modelo de demanda por água mineral e que, por alguma razão, não se observa a temperatura em cada região/mercado. Pelo lado da demanda, descontando outros efeitos como renda, em regiões com temperaturas mais elevadas serão consumidos relativamente mais litros de água (maior procura). Como consequência, pelo lado da oferta as firmas tenderão a elevar preços para reequilibrar o mercado. Portanto, movimentos não observados (erro) estão correlacionados com os preços.

Tal correlação é incompatível com o pressuposto básico do MQO e causa o problema conhecido tecnicamente como endogeneidade. Isto é, como o preço observado constitui equilíbrio de mercado, ele é também função dos choques que afetam a demanda. Os dados observados não são gerados apenas pela demanda. Constituem o resultado da interação entre oferta e demanda. Portanto, preços e quantidades são variáveis endógenas.

Ignorar o problema da endogeneidade tipicamente gera subestimação do coeficiente da variável preço em termos absolutos, o que resulta em subestimação das elasticidades (em módulo). Esta observação é extremamente relevante para a análise antitruste, pois quanto menor a sensibilidade do consumidor em relação a um aumento de preços, isto é, quanto menor a elasticidade (em módulo), maior é a capacidade de elevar preços por parte das firmas (maior poder de mercado). Portanto, ignorar a endogeneidade leva à conclusão enganosa de que as firmas possuem poder de mercado maior do que realmente detêm, o que facilitaria a ocorrência de falsos positivos (reprovação quando na realidade a conduta da firma ou o ato de concentração deveriam ser aprovados) no julgamento de casos antitruste.

Para solucionar o problema da endogeneidade é preciso aplicar métodos que utilizam variáveis instrumentais (VI). Assim como o MQO, tais métodos para modelos lineares são computacionalmente rápidos e se encontram disponíveis em todos os pacotes econométricos de grande circulação (e.g. Mínimos Quadrados em Dois

Estágios). No entanto, exigem do pesquisador a busca por instrumentos, isto é, variáveis correlacionadas com a variável endógena, mas não correlacionadas com o erro. Dito de outra forma, são variáveis que afetam a quantidade apenas pelo efeito sobre preços, não determinando a quantidade demandada diretamente. Utilizam-se deslocadores da oferta como instrumentos tais como salários e preços de insumos de produção.

Qualquer outra mudança exógena de preços, ou seja, que não seja influenciada pela quantidade demandada, tal como mudanças inesperadas de impostos, também podem ser utilizadas como deslocadores de oferta.

O uso de MQO justificar-se-ia apenas no caso de uma oferta perfeitamente elástica, em que choques na demanda não influenciam os preços de oferta. No entanto, este caso é irreal para a maioria dos mercados.

3. Modelos de Demanda por Produtos Diferenciados

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; a segunda classe inclui modelos (que serão analisados adiante) nos quais se assume que a escolha do produto se dá de forma indireta, a partir dos seus atributos.

Nos modelos do primeiro tipo, o pesquisador se depara com um 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 de cada bem rival r).

Em termos teóricos o número elevado de parâmetros não impõe severas limitações, o que não é verdadeiro para a 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*) – ver Deaton and Muellbauer (1980), também conhecido como NIDS (*Nearly Ideal Demand System*).

A causa do excesso de parâmetros se deve à hipótese de que os indivíduos auferem utilidade diretamente derivada do consumo dos bens produzidos no mercado. A solução para este problema, proposta por Lancaster (1966), consiste em assumir que os

consumidores atribuem utilidade às características dos bens e a demanda passa a depender das características destes bens (inclusive preço). 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 consumidores afinal demandam características e, em decorrência disso, é que demandam produtos. Ou seja, a demanda dos produtos é um reflexo da demanda das características. Uma versão mais sofisticada do modelo de Lancaster, e que se adequa de forma mais natural à análise empírica, se enquadra na classe de modelos de escolha discreta com utilidade aleatória (Random Utility Models, RUM). Estes modelos encontram vasta aplicação em Organização Industrial empírica. No entanto, iniciamos com um modelo de consumidor representativo para conjuntos fixos de produtos, a saber, o modelo Almost Ideal Demand System (AIDS/NIDS) citado anteriormente.

3.1. Modelo AIDS

O modelo AIDS, proposto por Deaton e Muellbauer (1980) é bastante popular na análise empírica da demanda em função da sua flexibilidade em representar sistemas de demanda por produtos diferenciados (Huse e Salvo, 2006).

De forma sucinta, a ideia básica é que o consumidor escolhe em etapas, alocando seus recursos primeiro entre o produto agregado (refrigerantes por exemplo) e outros bens substitutos⁷ e depois escolhendo o produto (nível mais desagregado) dentro do conjunto de bens que compõem o agregado (Coca Diet por exemplo).

No estágio superior modela-se a demanda pelo produto agregado.

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

Onde:

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

⁷ O modelo AIDS pode ser apresentado de forma mais genérica, com mais níveis de escolha. No entanto, neste texto apresenta-se a forma mais simples por questões puramente expositivas.

P_{nt} - índice de preços da indústria⁸

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

η_{nt} - Erro do estágio superior

No estágio inferior, modela-se a demanda por cada marca do produto relevante: o *share* de valor de cada marca j , s_{jt} é explicado por:

$$s_{jt} = \alpha_j + \beta_j \log(Y_{nt} / P_{nt}) + \sum_k \gamma_{jk} \log p_{kt} + Z_{nt} \theta_j + \varepsilon_{jt}$$

Onde

Y_{nt} - receita total da indústria

p_{jt} - preço do produto j .

s_{jt} - *share* de valor. É calculada pela razão entre receita da marca j (Y_{jt}) e receita da indústria (Y_{nt}), i.e., $s_{jt} = Y_{jt} / Y_{nt}$ ⁹

ε_{jt} - erro da equação

As elasticidades própria (η_{jj}) e cruzadas (η_{jk}) da demanda da marca j frente a aumentos no preço da marca k são dadas por

$$\eta_{jj} = \left[\frac{1}{s_j} (\gamma_{jj} - \beta_j s_j) + \left(1 + \frac{\beta_j}{s_j}\right) (1 + \gamma) s_j \right] - 1, \text{ (Elasticidade própria)}$$

$$\eta_{jk} = \frac{1}{s_j} (\gamma_{jk} - \beta_j s_k) + \left(1 + \frac{\beta_j}{s_j}\right) (1 + \gamma) s_k, \text{ (Elasticidade Cruzada)}$$

⁸ Utiliza-se frequentemente a aproximação linear do índice teórico, apresentada por Stone (1953), apud Huse e Salvo (2006), $\log P_{nt} = \sum_j s_j \log p_{jnt}$, onde s_j é o share de valor médio da marca j na região n .

Essa aproximação do índice teórico torna o modelo linear.

⁹ Define-se também o share médio, frequentemente usado como peso para cálculo do índice de Stone, isto é, $s_j = \sum_t Y_{jnt} / \sum_t Y_{nt}$

Note que no caso especial do primeiro estágio não ser estimado, supõe-se uma elasticidade-preço agregada unitária, isto é $\gamma = -1$, o que simplifica o cálculo das elasticidades. Além disso, as elasticidades podem variar no tempo, ao se substituir as *shares* médias do numerador das expressões de elasticidades (s_j na elasticidade própria e s_k nas elasticidades cruzadas), por *shares* em cada ano.

3.1.1. Estimação do modelo AIDS

A econometria de equações simultâneas pode ser usada para estimar a regressão acima. Como a variável dependente são as parcelas de mercado $s_{j,t}$, tais que $\sum_j s_{jt} = 1$, os erros das equações não são independentes entre si. O método mais imediato e de fácil implementação computacional é o SURE (*Seemingly Unrelated Regression*). No entanto, a exemplo da aplicação de MQO para produtos homogêneos, o preço é uma variável endógena devido à sua correlação com o erro. Neste caso o SURE gera estimadores inconsistentes.

A exemplo da demanda por produtos homogêneos, 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. A solução comumente adotada é o uso de métodos baseados em variáveis instrumentais. O método de Mínimos Quadrados em Três Estágios (MQ3E) é o mais utilizado para estimar o modelo AIDS.¹⁰

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 geograficamente distintos do mesmo produto. Os autores sugerem como instrumentos os preços do mesmo produto em

¹⁰ De modo simplificado, o método de MQ3E seria a aplicação de Variável Instrumental em sistemas de equações, como o SURE. Para mais detalhes, veja, por exemplo, Greene (2000). O Método Geral dos Momentos (MGM), que acomoda o MQ3E como um caso particular (Wooldridge, 2002), também pode ser utilizado.

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.

Outros tipos de instrumentos seriam os preços e os deslocadores de demanda defasados no tempo, sob certas hipóteses de comportamento temporal dos choques de demanda. Tanto no caso dos instrumentos sugeridos por Hausman et al., como no caso das variáveis defasadas, deve-se considerar a validade dos mesmos através de testes de especificação (testes de restrições de sobreidentificação, como Sargan e Hansen).

3.1.2. Vantagens e Desvantagens do Modelo AIDS

Uma das principais virtudes do modelo AIDS reside em sua flexibilidade, isto é, na 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, um total de $(N^2+3N-4)/2$ parâmetros devem ser estimados¹¹. Para um mercado com poucos produtos, este problema pode não ser grave. No entanto, para mercados com vários produtos (mercado de automóveis e cereais prontos para consumo, por exemplo) o modelo AIDS pode se tornar impraticável. De fato, para um mercado com 10 variedades, o modelo requer a estimação de 63 parâmetros, o que é manipulável econometricamente. No entanto, para um mercado com 60 variedades, número razoável para o mercado de automóveis, seria necessária a estimação de 1888 parâmetros.

3.2. Modelos de Escolha Discreta

¹¹ O número de parâmetros é obtido após a imposição das restrições de simetria e homogeneidade (Hausman et al., 2004).

Conforme a 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? 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 com a relação exponencialmente crescente entre o número de produtos e o número de parâmetros, permitindo aplicação em mercados caracterizados pela presença de muitas variedades (e.g. automóveis).¹²

3.2.1. Modelo Logit

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

$$u_{ij} = -\alpha p_j + x_j \beta + \xi_j + \varepsilon_{ij}$$

Onde:

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

ξ_j é um índice que agrupa outras características não incluídas no vetor x_j e

ε_{ij} adiciona à utilidade um erro estocástico de média zero¹³.

¹² Não discutiremos aqui a metodologia do Price Elasticity Model (PEM) pela pouca influência que o método teve nos vários casos em que foi empregada, apesar da grande importância junto a profissionais de marketing. A PEM é estruturada com base em entrevistas diretas com consumidores, em amostra representativa dos consumidores reais do produto, para os quais é apresentada uma série de painéis com combinações de preços e produtos, através dos quais se apura o impacto de variações de preços sobre as escolhas das marcas. A principal vantagem consiste em se mitigar os problemas de endogeneidade, visto que os preços são determinados de forma exógena. As principais críticas derivam de sua característica de pesquisa de opinião, quais sejam: i) presunção de igualdade de conhecimento entre os consumidores; ii) presunção de igualdade na distribuição e na apresentação dos produtos nos locais de compra; e iii) inobservância das restrições orçamentárias dos consumidores (afinal, dizer o que prefere em lugar de comprar o que prefere não depende da renda real).

¹³ O objetivo da divisão do vetor de características em x_j e ξ_j se tornará claro na seção empírica.

Por conveniência expositiva, é interessante reescrever a utilidade da seguinte forma:

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

onde

$$\delta_j = -\alpha p_j + x_j \beta + \xi_j$$

A especificação da demanda se completa com a definição do bem externo,¹⁴ 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 à resolução do seguinte problema de 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 ε_{ij} ¹⁵ é possível obter uma forma analítica para a probabilidade de o consumidor i escolher um determinado produto j , cuja fórmula é dada por

$$prob_{ij} = \frac{e^{\delta_j}}{\sum_{k=0}^1 e^{\delta_k}}$$

Observe que o lado direito da 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 $prob_{ij}$ à fatia de mercado do produto j (s_j), medida em quantidade ao invés de valor. Logo, a demanda por j é dada por

$$s_j = \frac{e^{\delta_j}}{\sum_{k=0}^N e^{\delta_k}}$$

¹⁴ O bem externo pode ser alguma alternativa que não faz concorrência direta entre os produtos em análise ou uma marca de referência.

¹⁵ Onde $f(\varepsilon_{ij}) = \exp(-\exp(\varepsilon_{ij}))$.

Ou, em sua forma log-linear

$$\ln s_j - \ln s_0 = \delta_j - \delta_0 = \delta_j = -\alpha p_j + x_j \beta + \xi_j$$

As elasticidades própria e cruzadas são dadas por:

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

$$\frac{p_r}{q_j} \frac{\partial q_j}{\partial p_r} = -\alpha p_r s_r \quad (\text{Elasticidade cruzada, (j diferente de r)})^{16}$$

3.2.1.1. 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 do mercado de automóveis novos, seria o número de potenciais compradores do produto automóvel 0Km;
- preços dos produtos p_j ;
- características dos produtos x_j ¹⁷.

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

O próximo passo consiste em estimar a seguinte equação:

$$\ln s_j - \ln s_0 = -\alpha p_j + x_j \beta + \xi_j$$

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

¹⁶ É possível demonstrar que as elasticidades preço-share são iguais às elasticidades preço-quantidade neste modelo, lembrando que $q_j = s_j M$, onde M = tamanho do mercado.

¹⁷ O modelo Logit pode ser calibrado para aproximar os dados, ao invés de ter seus parâmetros estimados. No caso da calibragem é possível prescindir deste conjunto de dados, no entanto isto não é verdade para a abordagem econométrica

exemplo da equação para produtos homogêneos, os preços são variáveis endógenas devido à correlação com o erro, índice que sumaria os atributos não incluídos no vetor x_j . Uma solução seria usar o método *VI*. Existem duas classes de variáveis instrumentais: *VI*'s construídas a partir das características dos produtos, propostas por Berry, Levinsohn e Pakes (1995), doravante denominado BLP, e *VI*'s construídas a partir da observação de preços do mesmo produto em mercados geograficamente distintos (Hausman et. al, 1994), como mencionados anteriormente.

As *VI*'s de BLP são:

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

As *VI*'s de Hausman são:

- as próprias características do produto;
- o preço do produto j no mercado A como instrumento para o preço do produto j no mercado B.

3.2.1.2. Problemas Conceituais do Modelo Logit

Mesmo após tratar todos os problemas econométricos que podem surgir em uma regressão e obter resultados convenientes, como demandas menos inelásticas do que seriam na aplicação de MQO, o modelo Logit pode não ser adequado para a análise de defesa da concorrência em muitos casos, pois apresenta sérios problemas conceituais, que se revelam claramente a partir do cálculo das elasticidades.

Observe a partir das elasticidades que o aumento da fatia de mercado do produto j decorrente do aumento percentual do preço p_r depende apenas do produto r . Isso 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 os mais próximos e mais distantes em relação ao grau de diferenciação, objetivo claramente comprometido

no modelo Logit.¹⁸ A restrição de que um aumento no preço do bem r afeta de forma idêntica todos os outros produtos de mercado é uma manifestação 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 um grau de substituição mais elevado. Outra solução, bem mais sofisticada se refere ao modelo Logit com coeficientes aleatórios (*Random Coefficients Model* ou *Mixed Logit Model*), onde a utilidade marginal pelos atributos características, incluindo preços, varia de consumidor para consumidor.

3.2.1.3. Sumário de Vantagens e Desvantagens do Logit

As principais vantagens do modelo Logit vêm de sua estrutura restritiva, ou seja, o pequeno número de parâmetros permite que o método seja empregado para estimação de mercados com presença de muitas variedades. Em adição, poucos dados são necessários, pois um *cross-section* seria suficiente para a estimação. Dados em painel podem ser empregados para melhorar a eficiência do modelo, mas não são indispensáveis como no modelo AIDS. Modelos lineares com variáveis instrumentais podem ser empregados, com instrumentos disponíveis a partir dos próprios dados empregados na estimação (Berry, Levinson e Pakes, 1995).

As desvantagens aparecem como o reverso das vantagens, ou seja, um padrão de substituição pouco razoável e definido *a priori*, com baixíssimo grau de flexibilidade e matriz de elasticidades com vários elementos idênticos, gerando resultados pouco plausíveis com relação à medição de poder de mercado. Assim, podemos dizer que o modelo Logit está distante do princípio *Let the data talk*. Está mais para *Let the Model Do Everything*.

¹⁸ Isto é muito importante para a delimitação do mercado relevante, pelo Teste do Monopolista Hipotético.

3.2.2. Logit Agrupado

Para tentar minimizar o problema de elasticidades-preço cruzadas similares entre produtos, sem gerar ainda um grande número de parâmetros, 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* multinomiais convencionais é corrigida parcialmente. Ela permanece válida para as alternativas existentes dentro dos subgrupos mais restritos, mas geralmente não o será entre os diferentes grupos e subgrupos.

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 contidas variedades com alto grau de substituição entre si. Estas classes podem ser diretamente representadas por determinadas 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), marca (Ford, Fiat, Peugeot, etc) e tipo de combustível (gasolina, álcool, etc), entre outros.

Nessa 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:

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

Assim temos a seguinte função utilidade indireta condicional do indivíduo i :

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

O primeiro termo aleatório ζ_{ig} representa o efeito de “choques” que afetam todos os produtos de um determinado grupo g e sua distribuição depende do parâmetro σ ($0 \leq \sigma < 1$)¹⁹. O segundo termo aleatório ε_{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 equação anterior como um modelo de coeficientes aleatórios ζ_{ig} , onde estes coeficientes operam apenas em variáveis *dummy* específicas.

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

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

Ainda de acordo com Berry (1994), assumindo uma forma funcional para a função utilidade e assumindo que v_{ij} tem uma distribuição de valor extremo²⁰ podemos obter uma forma analítica para a função de *market share*. De fato, é possível demonstrar que o *market share* do carro j de um determinado grupo g é dada por:

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

onde

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

O *market share* do grupo g é dado por

¹⁹ Conforme o valor de σ 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. A interpretação econômica é simples, quanto maior σ , mais segmentado é o mercado, ou seja, maior o grau de substituição entre bens do mesmo grupo do que entre bens pertencentes a grupos diferentes. No entanto, não se deve confundir segmento com mercado relevante. É possível ilustrar com teste do monopolista hipotético que o mercado relevante pode ser menor, igual ou maior que o conjunto de bens que compõe determinado segmento.

²⁰ Cardel mostra que se \mathcal{E} possui uma distribuição de valor extremo, então o termo $[\xi + (1 - \sigma)\varepsilon]$ também possui esta forma de distribuição [Cardell (1991) *apud* Berry (1994)].

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

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

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

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

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

onde s_0 é a proporção de consumidores que escolhem a alternativa externa (ou seja, não comprar um carro) e α , β , σ 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 ao Logit na flexibilidade das elasticidades-preço próprias e cruzadas. De fato, após manipulações algébricas, obtêm-se as seguintes fórmulas:

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

$$\frac{p_r \partial s_j}{s_j \partial p_r} = \frac{\alpha}{1-\sigma} p_r [(1-\sigma)s_r + \sigma s_{r/g}] \quad \text{Elasticidade cruzada (quando } j \text{ e } r \text{ pertencem ao mesmo grupo)}$$

$$\frac{p_r \partial s_j}{s_j \partial p_r} = \alpha p_r s_r \quad \text{Elasticidade cruzada (quando } j \text{ e } r \text{ pertencem a grupos diferentes)}$$

Ou seja, a elasticidade cruzada depende do fato de j e r estarem ou não em grupos diferentes.

3.2.2.1. Vantagens e Desvantagens do Logit Agrupado

O modelo Logit agrupado apresenta algumas vantagens em relação ao Logit original. 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 modelo LA é mais flexível do que o modelo Logit.

No entanto, na comparação com o modelo AIDS, o modelo LA ainda apresenta baixo grau de flexibilidade. Note que, dentro de cada segmento, as elasticidades cruzadas serão idênticas, replicando o problema do Logit. Ou seja, a matriz de elasticidades ainda apresenta alguns elementos idênticos.

Neste sentido o modelo LA ainda está distante do princípio *Let the data talk*, permanecendo mais para *Let the model do almost everything*.

3.2.3. Logit com Coeficientes Aleatórios ou Mixed Logit (Artigo BLP)

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

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

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),

ξ_j é um índice que agrupa outras características não incluídas no vetor x_j ²¹ e

ε_{ij} adiciona à utilidade um erro estocástico de média zero.

²¹ O objetivo da divisão do vetor de características em x_j e ξ_j se tornará claro na seção empírica.

α_i e β_{ik} são coeficientes aleatórios e seguem uma distribuição²² $\alpha_i = \alpha / v_{yi}$, onde o denominador corresponde à renda de cada consumidor cuja distribuição é log-normal por hipótese.

Assume-se também que β_{ik} segue uma distribuição normal padrão com média $\bar{\beta}_k$ e desvio padrão σ_k . Desta forma a utilidade pode ser reescrita da seguinte forma:

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

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 $x_j = (x_{j1}, x_{j2}, \dots, x_{jk}, \dots)$

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

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

Ao assumir uma distribuição de valor extremo do tipo II para ε_{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

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

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

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

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 do erro ξ_j impede o uso imediato

²² A exposição a seguir se aproxima daquela encontrada em Berry, Levinsohn and Pakes (1999).

das técnicas econométricas tradicionais. 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.

3.2.3.1. Prós e contras do modelo Mixed Logit (ML)

O modelo ML retém a principal vantagem dos modelos de escolha discreta i.e., um número reduzido de parâmetros. No entanto, ao contrário dos outros modelos desta classe apresentados anteriormente (Logit e Logit Agrupado), o modelo não impõe *a priori* um padrão de substituição entre os produtos, possuindo matriz de elasticidades que não contém forçosamente valores idênticos. Os instrumentos para controlar a endogenia seriam os mesmos indicados para o Logit. Por fim, em relação ao Logit e ao LA, o ML gera medidas mais plausíveis de poder de mercado.

Por outro lado, a grande desvantagem no uso do ML está na complexidade do modelo econométrico, que exige o uso de técnicas não usuais de regressão, que não estão disponíveis em pacotes econométricos. A carga computacional relativamente elevada do modelo, muitas vezes inviabiliza seu uso por técnicos de agências antitruste devido às restrições de tempo com que estes frequentemente se deparam.

3.3. Recomendações sobre o uso de Metodologias

Em relação a produtos homogêneos, a literatura se encontra bem consolidada de forma que não suscita comentários adicionais. No que diz respeito à demanda por produtos diferenciados, o modelo AIDS encontra vasta aplicação na estimação de modelos agregados de demanda. Possui boas propriedades (flexibilidade da matriz de elasticidades, possibilidade do uso de técnicas econométricas usuais e, conseqüentemente, fácil de implementação através de pacotes econométricos de grande circulação), constituindo desta forma 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 parâmetros²³.

Surge então a possibilidade de utilizar os modelos de escolha discreta (Logit, Logit Agrupado e Mixed Logit). Em adição, se os produtos apresentam características observadas que os diferenciam e são duráveis, os modelos de escolha discreta podem fazer mais sentido do que os modelos como o AIDS.²⁴ Dentre os quais os mais apropriados seriam o LA e o ML. O primeiro resolve parcialmente alguns problemas conceituais do Logit e possui a vantagem (em relação ao ML) de ser um modelo linear e, portanto, de fácil tratamento econométrico. No entanto, o LA exige do pesquisador justificativas para a escolha dos agrupamentos.

O ML, por sua vez, se mostra um modelo que possui excelentes propriedades (gera matriz de elasticidades flexível 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. Dessa forma, sua utilização deve ser reservada aos casos aos quais a agência de defesa da concorrência considera de elevada relevância e onde a análise de demanda tem o potencial de constituir elemento importante relevante para a análise antitruste.

Destaca-se que os modelos de escolha discreta exigem que as características dos produtos sejam conhecidas.

Quanto à disponibilidade de dados, o modelo AIDS exige dados ao longo do tempo ou a identificação de vários mercados geograficamente distintos para a identificação dos coeficientes. Séries longas devem ser usadas, ou pelo menos, com alta frequência (semanal), para aumentar o número de observações. Há sempre o *trade-off* de amostras temporais mais longas incluírem períodos onde o número, tipo e estratégias dos concorrentes e produtos ao longo do tempo variam, criando possíveis quebras estruturais nas estimativas.

²³ Mesmo nestes casos, é possível agregar produtos de forma a reduzir o número de parâmetros, o que pode ser útil na delimitação de mercado.

²⁴ Automóveis são um bom exemplo, em que o consumidor percebe e considera explicitamente as características dos produtos, em adição à marca. Um caso menos claro de produto diferenciado por características em adição à marca seriam refrigerantes (sabor frutas e *diet* ou não) e cereais. Apesar dos cereais terem conteúdos de açúcar e fibras variados, não parece tão claro que os consumidores decidam a compra principalmente por estas características.

Por outro lado, modelos de escolha discreta podem ser estimados com dados em *cross-section* (para uma só data) basicamente, pois o número de parâmetros não aumenta com o número de produtos envolvidos.

4. Síntese Crítica dos Estudos de Casos

No ano de 2009, o Departamento de Estudos Econômicos em conjunto com o Grupo Técnico de Métodos em Economia desenvolveu um trabalho com o objetivo de estudar a aplicação de métodos quantitativos na análise de demanda no âmbito da defesa da concorrência. Tal trabalho contemplou o estudo teórico dos principais modelos econométricos de estimação de demanda e uma revisão da aplicação de tais modelos em atos de concentração já julgados no Sistema Brasileiro de Defesa da Concorrência.

A parte teórica está sintetizada nas seções iniciais deste trabalho. Quanto à revisão do uso de modelos de estimação de demanda no âmbito do SBDC é necessário fazer algumas observações:

- i. os estudos pesquisados pertencem a atos de concentração já julgados pelo CADE;
- ii. a escolha dos atos de concentração pesquisados foi unicamente a existência nos autos de estudos de estimação de demanda;
- iii. outros pareceres econômicos versando sobre temas diferentes de estimação de demanda encontrados nos referidos autos não foram considerados neste trabalho;
- iv. as informações recolhidas dos pareceres estudados não foram quantificadas, pois, considerando o objetivo deste trabalho e o número limitado de pareceres pesquisados, tal detalhamento da informação não é relevante;
- v. no presente trabalho não se fará avaliação individualizada de nenhum estudo específico, nem à análise de determinado ato de concentração, tendo em vista que:
 - a. o objetivo deste trabalho é sugerir boas práticas na produção dos estudos entregues ao SBDC ou produzidos por seus membros;
 - b. existem, nos documentos pesquisados, informações confidenciais;
 - c. este documento é uma **versão pública**.

Os pareceres pesquisados foram apresentados por ocasião das análises dos seguintes atos de concentração no SBDC, os quais se referem, na maioria, a mercados de produtos diferenciados:

- **AC 08012.005846/1999-12** (Requerentes: Companhia Cervejaria Brahma e Companhia Antartica Paulista Indústria Brasileira de Bebidas e Conexos) – mercado de cervejas²⁵;
- **AC 08012.000212/2002-30** (Requerentes: Pepsico, Inc. e Companhia Brasileira de Bebidas) – mercado de bebidas isotônicas;
- **AC 08012.001697/2002-89** (Requerentes: Chocolates Garoto S/A e Nestlé Brasil Ltda. – mercado de chocolates;
- **AC 08012.007603/2003-66** (Requerentes: AGCO Corporation e Kone Corporation) – mercado de tratores;
- **AC 08012.000298/2007-13** (Requerentes: Recofarma Indústria do Amazonas Ltda., Spal Indústria Brasileira de Bebidas S/A e Sucos Del Valle do Brasil Ltda.) – mercado de sucos prontos para beber;
- **AC 08012.001383/2007-91** (Requerentes: Leão Júnior S/A Recofarma Indústria do Amazonas Ltda.) – mercado de chás prontos para beber;
- **AC 08012.003302/2007-97** (Requerentes: José de Sousa Cintra e Companhia de Bebidas das Américas – AmBev) – mercado de cervejas;
- **AC 08012.009107/2007-71** (Requerentes: DM Indústria Farmacêutica Ltda. e Hypermarchas S/A) – mercado de adoçantes industrializados de mesa de baixa caloria;
- **AC 08012.002813/2007-91** e **AC 08012.014599/2007-16** (Requerentes: Petróleo Brasileiro S/A e Braskem S/A) – mercado de polietilenos.

Na pesquisa realizada, observou-se que a teoria revista vem sendo aplicada em processos junto ao SBDC. Uma das razões pode ser a interação forte entre as empresas e a academia brasileira, com professores colaborando frequentemente para a elaboração de pareceres. Se os modelos iniciais eram de elasticidade constante, como no caso Antártica-Brahma, após o caso Nestlé-Garoto verifica-se uma mudança em direção a modelos do tipo NIDS/AIDS. Modelos do tipo Logit são relativamente pouco

²⁵ A indicação do mercado não representa a exata definição de mercado relevante adotada na análise do ato de concentração, mas apenas uma indicação de qual segmento foi objeto dos estudos de demanda analisados.

empregados, diante de suas qualidades positivas, e há apenas um caso de uso do Logit Agrupado, no caso AGCO-Kone.

Em alguns casos, a escolha do modelo ou do método de estimação foi justificada por sua simplicidade, tanto para a execução quanto para a compreensão dos resultados, contudo, essa vantagem deve ser comparada a possíveis desvantagens, principalmente com relação à adequação do modelo ao tipo de mercado e à consistência dos resultados. Observou-se em alguns pareceres que a utilização de técnicas simples ou teoricamente inadequadas (como a estimação por Mínimos Quadrados Ordinários²⁶) produziu resultados estatisticamente insignificantes.

Além disso, quando se obtém um grande número de resultados insignificantes, do ponto de vista estatístico, necessariamente deve-se investigar problemas de especificação do modelo. Entretanto, há casos em que o parecerista aparentemente ignora tais problemas, ou, pelo menos, não demonstra no seu estudo que investigou essa possibilidade ou que tenha realizado testes de robustez do seu modelo, passando simplesmente a interpretar uma série de números cuja falta de significância estatística torna sua credibilidade questionável.

Da mesma forma, ocorreu em alguns trabalhos que apresentaram resultados inesperados, contrários à expectativa inicial que se tem do mercado analisado (por exemplo, uma elasticidade própria com sinal positivo ou elasticidade cruzada com sinal negativo), uma tentativa de formular novas hipóteses sobre o mercado, como indicar que marcas supostamente concorrentes, sejam, na verdade, complementares. Não se pode negar que um modelo econométrico venha a apresentar resultados inesperados, evidenciando uma realidade de mercado diferente do senso comum, porém, nesses casos os cuidados para demonstrar a correta especificação do modelo devem ser ainda maiores, sob pena de que a ausência de testes estatísticos venha lançar dúvidas sobre a “nova realidade” apresentada.

Quanto às fontes dos dados, são relativamente escassas. Para produtos diferenciados, de consumo pessoal, a grande fonte de dados é a ACNielsen. Até poucos anos atrás, os dados eram médias bimestrais de diferentes canais. Mais recentemente, verifica-se o uso aprofundado de dados coletados junto a supermercados de modo

²⁶ Conforme salientado anteriormente (item 2), o MQO apresenta pressupostos teóricos que o tornam inadequado para estimações de demanda.

eletrônico (*check-out data*). Estes dados apresentam vantagens quanto à frequência e representatividade (erro de medida), mas algumas vezes com limitações regionais e de canal de venda. Para produtos diferenciados duráveis, as fontes são mais variadas, sendo possível obter dados junto às próprias requerentes, concorrentes e associações de classe.

A qualidade dos estudos é muito variada e os problemas apontados acima não são generalizados, mas devem ser ressaltados num trabalho como este que tem o objetivo de fomentar boas práticas. Se alguns estudos apresentam informações detalhadas sobre o modelo empregado, dados e testes de especificação, outros trazem informações dúbias ou incompletas sobre os modelos.

5. Comentários finais e Recomendações

A estimação de elasticidade-preço da demanda faz parte de qualquer análise sobre mercados e comportamento de firmas e consumidores. Seu uso em defesa da concorrência é amplo, auxiliando na determinação do mercado relevante pela metodologia do teste do monopolista hipotético e em modelos de simulação de fusões e conduta.

Este documento teve múltiplos objetivos. Primeiro, fazer uma resenha dos principais modelos de demanda por produtos na literatura internacional e empregados na prática antitruste no Brasil e no exterior, de modo similar a DeSouza (2009), Huse e Salvo (2006) e Hosken et al. (2002). Segundo, revisar o uso destes métodos nos casos apresentados junto ao SBDC. Por fim, fazer uma série de recomendações sobre estudos de demanda para o SBDC e seus usuários.

Para produtos diferenciados, há vários modelos a serem empregados, com diferentes complexidades. Em geral, percebe-se um *trade-off* entre flexibilidade dos modelos e número de observações necessárias para estimar os coeficientes de modo preciso. Enquanto o modelo NIDS/AIDS aparece como o usual para produtos diferenciados, as estimativas podem ser bastante imprecisas pela quantidade de parâmetros a serem estimados. Casos especiais do modelo NIDS podem ser usados, como o de elasticidades constantes (log-log), ao custo de imposição de restrições sobre o valor das elasticidades.

Para produtos diferenciados com características observáveis percebidas pelos consumidores, modelos de escolha discreta como o Logit, Logit Agrupado e Logit com

Coefficientes Aleatórios (Mixed Logit) são o caminho natural a seguir. Estes modelos apresentam outro *trade-off*, de facilidade de estimação e flexibilidade. O caso extremo seria o Logit, que é um modelo com restrições fortes sobre as elasticidades-preço cruzadas. Já o modelo Mixed Logit supera tais restrições ao custo de exigir habilidades de programação avançada do usuário em softwares estatísticos e muito tempo de processamento e estimação.

Não é demais lembrar que estes *trade-offs* podem ser particularmente agudos junto ao SBDC, em que a tempestividade, parcimônia e facilidade de interpretação são fundamentais para a tomada de decisão.

Em todos os casos deve-se ter uma preocupação com a determinação conjunta de preços e quantidades, ou seja, com a endogeneidade de preços. A violação da hipótese de métodos de estimação usuais como mínimos quadrados ou de escolha discreta leva à subestimação das elasticidades de demanda (mais próximas de zero do que realmente são), dando a falsa impressão de pouca substitutibilidade entre os produtos considerados. Isto pode levar a mercados relevantes mais restritos do que realmente são.

O problema de endogeneidade de preços não é de solução fácil, embora seja conhecido. O uso de variáveis instrumentais, recomendado neste caso, esbarra muitas vezes na falta de variáveis que desloquem a curva de oferta, para que a demanda seja identificada pelos métodos econométricos. Há alguns instrumentos conhecidos, como custos, características dos produtos concorrentes e preços em outras regiões - estes dois últimos em geral associados a autores como Berry, Levinsohn e Pakes (1995), e Hausman et al. (1994), respectivamente. Todavia, em cada caso deve-se justificar o uso do conjunto de instrumentos com argumentos puramente econômicos e, em etapa posterior, avaliar econometricamente sua qualidade (forte correlação com preços) e validade (exogeneidade em relação à quantidade), através de testes de especificação (descritos em qualquer bom livro de econometria moderno).

Em termos gerais, deve-se destacar que a experiência de métodos de estimação de demanda junto a processos no SBDC está aumentando e melhorando em qualidade. Todavia, a qualidade dos estudos é muito variada. Se alguns estudos apresentam informações detalhadas sobre o modelo empregado, dados e testes de especificação, outros trazem informações dúbias ou incompletas sobre os modelos. Não se deve esquecer que todo modelo econométrico apresenta propriedades estatísticas desejáveis – ou seja, credibilidade em seus resultados – sob algumas hipóteses. Estas hipóteses

podem e devem ser verificadas, como por exemplo, o citado problema de endogeneidade de variáveis explicativas (preços).

Por fim, este documento também teve como um dos seus objetivos principais estabelecer uma série de recomendações com respeito ao conteúdo de pareceres econômicos sobre demanda de cunho quantitativo submetidos aos processos a serem apreciados pelo CADE. Após rever uma série de pareceres anexados a vários processos analisados anteriormente pelo Conselho, conclui-se pela elaboração de algumas recomendações e observações que seguem..

Sobre os dados, chamamos atenção para a importância de o pesquisador sempre explicitar a origem, as técnicas de tratamento (agregações, deflacionamentos, dessazonalização, por exemplo) e as limitações espaciais e temporais dos dados utilizados na estimação das elasticidades. Se o trabalho contém várias estimações com diferentes técnicas, é importante manter a mesma base de dados para efeitos comparativos. O ideal é que estudos dessa natureza venham sempre acompanhados dos respectivos dados (inclusive em meio eletrônico) utilizados nas regressões, bem como que seja indicado o pacote estatístico ou econométrico no qual as estimações foram realizadas. Isto permitirá à autoridade de antitruste entender melhor as estimativas e o processo de análise levado a cabo e, se necessário, replicar os testes com análise de robustez para chegar a conclusões críveis.

Entendemos como imprescindível que os estudos de estimação de demanda instruídos em processo de atos de concentração apresentem, juntamente com os resultados, os valores dos testes estatísticos consagrados na literatura econométrica, a saber: significância estatística “t de *student*”, heterocedasticidade, autocorrelação residual, raiz unitária e co-integração (séries temporais). Deve-se definir claramente a forma de cálculo das elasticidades, bem como dos respectivos desvios-padrão, indispensáveis aos testes de significância. Também é relevante definir e manter o mesmo nível de significância ao longo de toda a análise, onde talvez caiba à própria autoridade determinar os níveis de confiança adequados.

Quanto às técnicas utilizadas, desde que haja alternativas, recomendamos que a estimação das elasticidades empregue mais de um método. Em verdade, o mais interessante é, na medida do possível, comparar as elasticidades obtidas por meio de diferentes técnicas e especificações, de forma que os resultados possam ser ratificados, ou, o que é igualmente salutar, questionados quanto à sua validade.

Para tornar as análises mais ricas e as simulações mais próximas da realidade, é interessante que a estimação da demanda considere a existência de bens substitutos, de diferentes marcas e formas de apresentação dos produtos, bem como contemple esquemas de tomada de decisão por parte dos consumidores.

Ademais, e como já frisado anteriormente, em se tratando de estimação de demanda, os estudos devem sempre levar em conta e propor soluções viáveis para equacionar os problemas de endogeneidade, onde cabe explicitar exatamente quais variáveis instrumentais foram utilizadas para especificar a demanda e quais as justificativas econômicas para sua escolha.

Deve-se deixar claro que a econometria servirá para mensurar e estabelecer margens de erros para as estimativas realizadas. O estudo econométrico não necessariamente servirá para uma conclusão categórica a respeito da definição do mercado relevante ou demais questões de análise antitruste que o analista venha a se defrontar, sendo parte importante da análise.

Acreditamos que a credibilidade e importância dos estudos de demanda aumentem ao longo do tempo, com a popularização dos métodos e cuidados com suas aplicações. Embora não seja um fim em si, nem parte única e com supremacia frente a outras evidências em processos junto ao SBDC, a estimação de funções de demanda podem ser de grande valia para a compreensão da realidade e no auxílio na tomada de decisões.

6. Bibliografia

- BERRY, S. (1994). **Estimating Discrete-Choice Models of Product Differentiation.** *Rand Journal*, 25(2), pp. 242-262.
- BERRY, S.; LEVINSOHN, J.; 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.; MUELLBAUER, J. (1980). **An Almost Ideal Demand System.** *American Economic Review*, 70, 312-326.
- DeSOUZA, S. (2009). **Análise Empírica de Modelos Agregados de Demanda.** Working Paper, CAEN, UFC.
- _____, S. (2009). **Antitrust Mixed Logit Model.** Mimeo, CAEN, UFC.
- GUJARATI, D. (2006). **Econometria Básica.** Rio de Janeiro, CAMPUS, 4.ed.
- HAUSMAN, J; LEONARD, G.; ZONA J. (1994) “Competitive Analysis with Differentiated Products” *Annales d’Economie et de Statistique*, 34, 159-180
- HOSKEN, D. et al. (2002). “Demand System Estimation and its Application to Horizontal Merger Analysis”. *Mimeo*, Federal Trade Commission, EUA.
- HUSE, C. e SALVO, A. (2006). **Estimação e identificação de Demanda e de Oferta.** In: *Métodos Quantitativos em Defesa da Concorrência e Regulação Econômica*, Eds. Eduardo P. S. Fiuza e 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*.
- 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.