

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

PATRICIA GONZALEZ CONTE

**MODELOS DE DEMANDA: UMA
PROPOSTA METODOLÓGICA
PARA COMPARAÇÃO DOS
MODELOS MAIS UTILIZADOS
PELOS ÓRGÃOS DE DEFESA DA
CONCORRÊNCIA BRASILEIROS E
APLICAÇÃO EMPÍRICA.**

Rio de Janeiro
2017

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

PATRICIA GONZALEZ CONTE

**MODELOS DE DEMANDA: UMA
PROPOSTA METODOLÓGICA
PARA COMPARAÇÃO DOS
MODELOS MAIS UTILIZADOS
PELOS ÓRGÃOS DE DEFESA DA
CONCORRÊNCIA BRASILEIROS E
APLICAÇÃO EMPÍRICA.**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Economia da Indústria e Tecnologia, Instituto de Economia, Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Mestre em Economia.

Orientador: Eduardo Pontual Ribeiro

Rio de Janeiro
2017

FICHA CATALOGRÁFICA

C761 Conte, Patricia Gonzalez.

Modelos de demanda: uma proposta metodológica para comparação dos modelos mais utilizados pelos órgãos de defesa da concorrência brasileiros e aplicação empírica / Patricia Gonzalez Conte. - 2017.

66 p. ; 31 cm.

Orientador: Eduardo Pontual Ribeiro.

Dissertação (mestrado) – Universidade Federal do Rio de Janeiro, Instituto de Economia, Programa de Pós-Graduação em Economia da Indústria e Tecnologia, 2017.

Bibliografia: f. 60 – 64.

1. Econometria. 2. Demanda (modelos). 3. Defesa da concorrência. 4. Concentração de mercado. I. Ribeiro, Eduardo Pontual, orient. II. Universidade Federal do Rio de Janeiro. Instituto de Economia. III. Título.

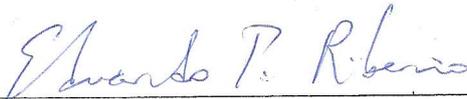
CDD 330.015195

PATRICIA GONZALEZ CONTE

**Modelos de Demanda: uma proposta metodológica para
comparação dos modelos mais utilizados pelos órgãos de
defesa da concorrência brasileiros e aplicação empírica.**

Dissertação de Mestrado apresentada ao Programa
de Pós-Graduação em Economia da Indústria e
Tecnologia, Instituto de Economia, Universidade
Federal do Rio de Janeiro, como requisito parcial
à obtenção do título de Mestre em Economia..

Aprovado em



Prof. Dr. Eduardo Pontual Ribeiro (Orientador)
Instituto de Economia/UFRJ



Prof. Dr. Getúlio Borges da Silveira Filho

Instituto de Economia/UFRJ



Prof.Dr. Eduardo Pedral Sampaio Fiuza
Instituto de Pesquisa Econômica Aplicada / IPEA

Rio de Janeiro
2017

A minha família.

AGRADECIMENTOS

Agradeço especialmente ao meu orientador, Eduardo Pontual Ribeiro, pela paciência e dedicação.

Aos professores Getúlio Borges da Silveira Filho e Eduardo Pedral Sampaio Fiuza pelas críticas e sugestões ao trabalho.

Ao Instituto de Economia da UFRJ pela oportunidade de realizar o mestrado.

E, por fim, agradeço aos amigos e a minha família por todo apoio e ajuda.

RESUMO

Conte, Patricia Gonzalez. **Modelos de Demanda: uma proposta metodológica para comparação dos modelos mais utilizados pelos órgãos de defesa da concorrência brasileiros e aplicação empírica..** 2017. 66 f. Dissertação (Mestrado em Economia da Indústria e Tecnologia) – Instituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017.

Na literatura existem vários modelos de demanda propostos que podem dar resultados muito diferentes em relação às variáveis-chaves para análise setorial e de concorrência, como as elasticidades. Temos também o problema de que nos casos em que os modelos que estão sendo comparados não possuem uma relação de aninhamento, não é possível utilizar a econometria básica para selecionar o modelo que melhor explicaria a curva de demanda. Alguns trabalhos já se propuseram a realizar esse tipo de comparação, mas ou eles envolviam modelos de demanda de fatores ou utilizavam testes com maiores dificuldades práticas. O trabalho então propõe uma metodologia estatística para comparar modelos de demanda; sendo eles aninhados, ou não, e com as variáveis dependentes de cada modelo diferentes ou iguais. Um teste do tipo de Wald será usado para os modelos aninhados e a abordagem de “modelo artificial” ou “aninhamento artificial” para os modelos não-aninhados. Dentro da segunda abordagem, utilizaremos, mais especificamente o teste PE e o teste BEM. Essa metodologia é adaptada para os principais modelos utilizados pelos órgãos de defesa da concorrência no Brasil: log-linear, LAIDS, PCAIDS e logit. Para ilustrar os métodos apresentados, uma aplicação empírica é realizada a partir dos dados semanais de uma rede de supermercado americana. O resultado da comparação pareada foi de que o modelo LAIDS seria superior ao modelo log-linear e ao PCAIDS, mas que não seria possível fazer qualquer afirmação relacionada ao modelo Logit.

Palavras-chave: Modelos de demanda, Testes de hipóteses para modelos não aninhados, Defesa da concorrência, Atos de concentração, Simulação de fusão.

ABSTRACT

Conte, Patricia Gonzalez. **Modelos de Demanda: uma proposta metodológica para comparação dos modelos mais utilizados pelos órgãos de defesa da concorrência brasileiros e aplicação empírica..** 2017. 66 f. Dissertação (Mestrado em Economia da Indústria e Tecnologia) – Instituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017.

In the literature there are several proposed demand models that can give very different results in relation to the key variables for sector analysis and competition, such as elasticities. We also have the problem that in cases where the models being compared do not have a nesting relationship, it is not possible to use basic econometrics to select the model that would best explain the demand curve. Some studies have already proposed to make this type of comparison, but either they involved factor demand models or used tests with greater practical difficulties. The work then proposes a statistical methodology to compare models of demand; Being they nested, or not, and with the dependent variables of each model different or equal. A wald-type test will be used for nested models and the "artificial model" or "artificial nesting" approach for non-nested models. Within the second approach, we will use, more specifically the PE test and the BEM test. This methodology is adapted to the main models used by the Brazilian defense agencies: log-linear, LAIDS, PCAIDS and logit. To illustrate the methods presented, an empirical application is made from weekly data from an American supermarket chain. The result of the paired comparison was that the LAIDS model would be superior to the log-linear model and the PCAIDS, but that it would not be possible to make any assertion related to the Logit model.

Keywords: Demand models, Hypothesis tests for non-nested models, Defense of competition, Acts of concentration, Fusion simulation.

LISTA DE TABELAS

Tabela 3.1: Comparação de modelos	35
Tabela 4.1: Estatísticas das variáveis utilizadas	38
Tabela 5.1: Modelo log-linear estimado por SUR	43
Tabela 5.2: Modelo log-linear estimado por MQ3E	44
Tabela 5.3: Modelo LAIDS estimado por SUR	45
Tabela 5.4: Modelo LAIDS estimado por MQ3E	46
Tabela 5.5: Modelo PCAIDS estimado por MQO	49
Tabela 5.6: Modelo PCAIDS estimado por MQ2E	50
Tabela 5.7: Modelo PCAIDS - Coeficientes obtidos a partir da estimação de γ_{11} por MQO.	50
Tabela 5.8: Modelo PCAIDS - Coeficientes obtidos a partir da estimação de γ_{11} por MQ2E.	50
Tabela 5.9: Modelo Logit estimado por SUR	51
Tabela 5.10: Modelo Logit estimado por MQ3E	52
Tabela 5.11: Elasticidades calculadas a partir dos coeficientes obtidos por SUR.	53
Tabela 5.12: Elasticidades calculadas a partir dos coeficientes obtidos por MQ3E.	54
Tabela 5.13: Testes das restrições do modelo PCAIDS a partir dos coeficientes encontrados pela estimação por SUR.	55
Tabela 5.14: Testes das restrições do modelo PCAIDS a partir dos coeficientes encontrados pela estimação por MQ3E.	56
Tabela 5.15: Testes PE e BEM sem auxílio de variável instrumental.	56
Tabela 5.16: Testes PE e BEM com auxílio de variável instrumental	57

LISTA DE ABREVIATURAS E SIGLAS

AIDS	Almost Ideal Demand System
CADE	Conselho Administrativo de Defesa Econômica
FFF	Flexible Functional Forms Factor Demand System
GLT	Generalized Leontief Factor Demand System
KLIC	Kullback - Leibler Information Criterion
LAIDS	Linear Almost Ideal Demand System
MMG	Método Geral dos Momentos Generalizado
MQO	Mínimos Quadrados Ordinários
MQ2E	Mínimos Quadrados em Dois estágios
MQ3E	Mínimos Quadrados em Três Estágios
NHQDR	Quadratic Non-Homogeneous Factor Demand System
QDR	Quadratic Factor Demand System
TLG	Translog Factor Demand System
SGM	Symmetric Generalized McFadden Factor Demand System
SUR	Seemingly Unrelated Regressions

SUMÁRIO

1	INTRODUÇÃO	12
2	MODELOS DE DEMANDA	15
2.1	Modelo Log-linear	16
2.2	Modelo LAIDS	16
2.3	Modelo PCAIDS	19
2.4	Modelo Logit	20
3	MÉTODOS E AJUSTES PARA A COMPARAÇÃO DOS MODELOS	22
3.1	Os Testes de Hipóteses	22
3.2	A implementação dos Testes para Modelos de Demanda	30
3.2.1	Comparação do Modelo Log-Linear e Logit	30
3.2.2	Comparação do Modelo Log-linear e LAIDS ou PCAIDS	31
3.2.3	Comparação Modelo Logit e LAIDS ou PCAIDS	33
4	DADOS	36
5	APLICAÇÃO EMPÍRICA	39
5.1	Métodos de Estimação dos Modelos	39
5.2	Estimação dos modelos	42
5.2.1	Modelo Log-linear	42
5.2.2	Modelo LAIDS	44
5.2.3	Modelo PCAIDS	46
5.2.4	Modelo Logit	51
5.3	Elasticidades	52
5.4	Resultados dos testes	55
5.4.1	Modelos Aninhados	55
5.4.2	Modelos Não Aninhados	56
	CONCLUSÃO	58
	REFERÊNCIAS	60
	APÊNDICE A TESTE DE ENDOGENEIDADE	65
	APÊNDICE B TESTE DOS INSTRUMENTOS	66

1 INTRODUÇÃO

O objetivo do trabalho é propor uma metodologia estatística para comparar as principais formas funcionais de demanda utilizadas pelos órgãos de defesa da concorrência no Brasil. A metodologia será aplicada em um conjunto de dados a fim de demonstrar sua aplicabilidade e indicar qual dos modelos selecionados melhor explicaria a curva de demanda.

Na área de estudos econômicos, o surgimento de diferentes teorias para explicar um mesmo fenômeno gera, muitas vezes, modelos estatísticos também diferentes. Em decorrência aos possíveis questionamentos sobre qual deles representaria melhor o fenômeno em estudo, foram desenvolvidos testes que tratam do problema de seleção dos modelos. Uma identificação necessária para aplicação desses testes é se os modelos em questão são aninhados, ou não (Pesaran e Ulloa, 2006).

Modelos são considerados aninhados se um deles for o caso especial do outro, ou seja, caso seja possível chegar de um modelo a outro através da imposição de restrições sobre os parâmetros. Na realização do teste de hipóteses, quando a hipótese nula estiver aninhada à alternativa, podemos utilizar procedimentos clássicos baseados em razão de verossimilhança, Wald ou no multiplicador de Lagrange. Caso contrário os testes clássicos não poderão ser aplicados diretamente; sendo necessário então utilizar testes específicos da literatura de “testes não aninhados” (Pesaran e Ulloa, 2006).

Em análise de regressão, há três grupos de métodos para análise de modelos não aninhados: (1) o teste Cox (Cox, 1961 e 1962), com extensão para modelos de regressão lineares proposta por Pesaran (1974) e equações simultâneas não lineares

por Pesaran e Deaton(1978); (2) o método baseado na construção de um modelo geral artificial, explorado por Atkinson (1970), Davidson e Mackinnon(1981,1982) e Bera e McAleer(1989), que cria um nova hipótese na qual os modelos em questão são aninhados; (3) e a linha "encompassing test", originalmente considerado por Deaton (1982) e Dastoor(1983) e mais à frente desenvolvido por Mizon e Richard(1986), (Pesaran e Ulloa, 2006).¹

Na economia aplicada, essa variedade de modelos ocorre em especial acerca dos modelos de demanda, usados em estudos sobre produção de bens e consumo para tentar descrever o comportamento de mercados e consumidores. Na área de estudo sobre fusões horizontais, os modelos são utilizados na obtenção das elasticidades-preço da demanda própria e cruzada. As elasticidades são necessárias para realizar a delimitação dos mercados relevantes em simulações de fusão. A partir da simulação pode-se mensurar os possíveis efeitos sobre preços e quantidades decorrentes desse ato de concentração (Bomfim, 2016), ou indicadores de rivalidade como UPP. Diferentes modelos empíricos de demanda produzem diferentes elasticidades estimadas para um mesmo ponto e diferentes variações de elasticidade e preço em decorrência das formas e inclinações das curvas de demanda (Crooke et al, 1999). Essas disparidades em relação às elasticidades pode ocasionar resultados divergentes sobre a simulação, dependendo de qual dessas elasticidades foi utilizada, e, assim, resultados divergentes sobre a aprovação, ou não, de um ato de concentração (Bomfim, 2016). No Brasil, a escolha do modelo de demanda pelos órgãos de defesa da concorrência é

¹Outro método relevante, mas que não segue a linha de teste de hipóteses, para comparação de modelos não aninhados é o "Teste de seleção de modelo de Vuong" (Vuong's 1989), apud (Pesaran e Ulloa, 2006), que utiliza o *Kullback -Leibler Information Criterion* (KLIC). Não há na literatura um consenso se a seleção de modelos deve ser feita pela qualidade do ajuste - ou seja, pelo critério de informação - ou teste de hipóteses. Davidson e MAckinnon, por exemplo, afirmam que se o objetivo é escolher o melhor modelo dentro de um conjunto, não importando se ele é falso ou não, devemos usar um procedimento próprio para isso, que geralmente envolve algum tipo de critério de seleção. O método de modelos artificiais seria então direcionado para a especificação do modelo (Davidson e MAckinnon, 2004). Contudo, Bera and McAleer (1989), apud Manera (1991), não concordam com essa colocação pois, para estes autores, os testes não aninhados seriam capazes de uma afirmação probabilística a respeito da seleção de modelos.

arbitrária: não há muitos estudos sobre e, portanto, nem indicação sobre qual delas seria mais apropriada.

O trabalho então se propõe a comparar os modelos de demanda mais utilizados pelo órgão de defesa da concorrência no Brasil (DEE, 2009): o log-linear, *Linear Almost Ideal Demand System* (LAIDS) e Logit. Também adicionamos o *Proportionally Calibrated Almost Ideal Demand System* (PCAIDS), que foi utilizado na análise de alguns atos de concentração relevantes do Conselho Administrativo de Defesa Econômica (CADE).² Nesse trabalho utilizaremos testes de hipóteses para comparar os modelos. O teste F será usado para os modelos aninhados e a abordagem de “modelo artificial” ou “aninhamento artificial” para os modelos não-aninhados. Dentro da segunda abordagem, utilizaremos, mais especificamente, os testes propostos por Davidson e Mackinnon(1981), Davidson, White e Mackinnon(1983) e Bera and McAleer (1983). O principal atrativo dessa abordagem é a simplicidade na hora da aplicação dos teste ao mesmo tempo que, segundo Davidson e Mackinnon(1993), seria assintoticamente equivalente ao teste Cox:

“[...] o teste Cox e seus derivados, que são assintoticamente equivalentes ao teste J e tem igualmente propriedades de amostra finita limitadas, é mais complicado de computar do que o teste J e, portanto, não parece haver razões para pôr-lo em prática.” (Davidson e Mackinnon, 1993)

Além da introdução, o trabalho conta com quatro capítulos e uma conclusão. No capítulo 1, são apresentados os modelos de demanda. O 2 é destinado aos testes que tratam da comparação de modelos aninhados, não aninhados com mesma variável dependente e não aninhado com variáveis dependentes diferentes. No capítulo 3, são apresentados os dados que serão utilizados para a aplicação empírica. No 4, os sistemas de demanda são estimados, calculadas as elasticidades e apresentados os resultados dos testes estatísticos. Por fim, temos as conclusões.

²Atos de Concentração nº 08012.001697/2002-8 e 08700.009988/2014-09

2 MODELOS DE DEMANDA

Nesse capítulo, nos basearemos no artigo DEE (2009). A equação de demanda é uma expressão matemática que relaciona a quantidade demandada de um bem e fatores que afetam a iniciativa do consumidor em comprar esse bem, como preço ou renda. A equação de demanda assume, então, a seguinte forma geral:

$$G(Q, P, \epsilon, D, \Omega) = 0 \quad (2.1)$$

Onde:

P - vetor de preço.

Q - vetor de quantidade.

D - vetor de deslocadores da demanda (Como renda, características do produto...).

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

Ω - vetor de parâmetros do modelo

A existência de variadas formas funcionais para demanda decorre dos diferentes modelos de preferência.

2.1 Modelo Log-linear

Em um mercado caracterizado por produtos heterogêneos, o modelo log-linear poderá então ser escrito da seguinte forma para um produto j em um mercado com N produtos:

$$\log Q_j = \alpha_{j0} + \sum_{k=1}^N \alpha_{jk} \log P_k + \alpha_{jD} \log D + \epsilon_j \quad (2.2)$$

As elasticidades própria e cruzada do modelo log-linear serão, sendo j diferente de k , iguais a

$$\eta_{jj} = \frac{\Delta Q_j}{Q_j} \frac{P_j}{\Delta P_j} = \alpha_{jj} \quad (2.3)$$

$$\eta_{jk} = \frac{\Delta Q_j}{Q_j} \frac{P_k}{\Delta P_k} = \alpha_{jk} \quad (2.4)$$

Como podemos observar, as elasticidades do modelo serão constantes. Essa é um hipótese pouco plausível e deve ter seu uso, portanto, justificado (DEE, 2009).

2.2 Modelo LAIDS

Proposto por Deaton e Muellbauer (1980), o modelo AIDS ganhou bastante popularidade devido a sua flexibilidade em representar sistemas de demanda por

produtos diferenciados (Huse; Salvo, 2006). O modelo funciona a partir da ideia de níveis: o consumidor escolhe alocar seus recursos primeiro em um produto agregado (refrigerante de cola, por exemplo) e outros bens substitutos, e depois escolhendo um produto dentre os bens que compõe o agregado (uma marca de refrigerante em particular, por exemplo).

No estágio superior modela-se a demanda pelo produto agregado através do modelo log-linear:

$$\log Q = \beta_0 + \gamma \log I + D\Theta + \epsilon \quad (2.5)$$

Onde:

I - índice de preços do setor. ¹

ϵ - erro do estágio superior.

No nível inferior, a demanda será referente a cada marca do produto agregado. O share de valor de cada marca j , s_j , é explicado por:

$$s_j = \beta_{0j} + \beta_j \log \left(\frac{Y}{I} \right) + \sum_k^N \gamma_{jk} \log P_k + D\Theta_j + \epsilon_j \quad (2.6)$$

¹Huse e Salvo (2006) ressaltam em seu artigo que a estimação através desse índice pode ser trabalhosa e pode haver problemas de falta de convergência decorrente da sua não linearidade. Eles propõem como alternativa a utilização do índice de Stone, $\log I = \sum_j^N s_j \log P_j$, para trabalhos empíricos. Essa aproximação do índice teórico torna o modelo AIDS no *Linear Almost Ideal Demand System* (LAIDS) (DEE 2009).

Onde:

Y - gasto total nos produtos

P - preço do produto j .

s_j - share de valor. É calculada pela razão entre gasto da marca j (Y_j) e gasto da indústria (Y).

ϵ_j - 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:

$$\eta_{jj} = \left[\frac{1}{s_j}(\gamma_{jj} - \beta_j s_j) + \left(1 + \frac{\beta_j}{s_j}\right) (1 + \gamma) s_j \right] - 1 \quad (2.7)$$

E

$$\eta_{jk} = \frac{1}{s_j}(\gamma_{jk} - \beta_j s_k) + \left(1 + \frac{\beta_j}{s_j}\right) (1 + \gamma) s_k \quad (2.8)$$

Temos como principal vantagem do modelo LAIDS quando comparado aos modelo log-linear e, como veremos, ao logit, um maior flexibilidade, permitindo qualquer padrão para a matriz de elasticidades. Em contrapartida, é o modelo com maior número de parâmetros a serem estimados.

2.3 Modelo PCAIDS

O PCAIDS foi apresentado por Epstein e Rubinfeld (2002) como um modelo derivado do AIDS, conservando assim muitas de suas propriedades, mas com a capacidade de ser aplicado a casos onde haja limitação de dados ou problemas de estimação. Para calcular as elasticidades a partir desse modelo, são necessários apenas os market shares, a elasticidade preço da indústria e de um produto (Epstein e Rubinfeld, 2002).

O modelo PCAIDS parte da premissa de proporcionalidade, ou seja, a quantidade de vendas perdida por um produto devido a um aumento de preço seria redistribuída às demais proporcionalmente de aos seus market shares. A equação de share de valor para o PCAIDS pode ser escrita da seguinte forma:

$$s_j = \beta_{0j} + \sum_k^N \gamma_{jk} \log p_k + D\Theta_j + \epsilon_j \quad (2.9)$$

De acordo com Epstein e Rubinfeld (2001), a partir dos market shares e de um γ_{kk} conhecido é possível obter os demais γ_{jj} e γ_{jk} pelas seguintes equações:

$$\gamma_{jj} = \frac{s_j}{1 - s_k} \frac{1 - s_j}{s_k} \gamma_{kk} \quad (2.10)$$

$$\gamma_{jk} = \frac{-s_k}{(1 - s_j)} \gamma_{jj} \quad (2.11)$$

Além disso, no modelo PCAIDS, a renda, e portanto a variável relativa ao índice de preços (PI), não tem efeito significativo sobre o market share de cada marca (Bomfim, 2016). Dessa forma, temos $\beta_j = 0$ na equação (2.6). Assim, chegaremos às seguintes elasticidades:

Elasticidade própria

$$\eta_{jj} = \left[\frac{1}{s_j} \gamma_{jj} + (1 + \gamma) s_j \right] - 1 \quad (2.12)$$

Elasticidade cruzada

$$\eta_{jk} = \frac{1}{s_j} \gamma_{jk} + (1 + \gamma) s_k \quad (2.13)$$

O modelo PCAIDS nos oferece, como grande vantagem, precisarmos estimar apenas um parte do sistema de demanda.

2.4 Modelo Logit

O modelo Logit assume que os consumidores atribuem utilidade às características dos bens dos quais a demanda passa a depender - inclusive preço (DEE 2009). Sendo w_j o share de quantidade, D_j um vetor de características dos produtos de dimensão N , ξ_j o índice que agrupa outras características não incluídas no vetor D_j , ϵ_j - adiciona à utilidade um erro estocástico de média zero e $j = 0$ o produto externo, temos, segundo DEE (2009), a seguinte equação para ser estimada:

$$\ln w_j - \ln w_0 = -\pi P_j + D_j \omega + \xi_j + \epsilon_j \quad (2.14)$$

A elasticidade própria é dada então por

$$\eta_{jj} = \frac{P_j}{Q_j} \frac{\Delta P_j}{\Delta Q_j} = -\pi P_j [1 - w_j] \quad (2.15)$$

E a cruzada, onde j é diferente de k,

$$\eta_{jk} = \frac{P_j}{Q_j} \frac{\Delta P_j}{\Delta Q_k} = \pi P_k [w_k] \quad (2.16)$$

O modelo Logit permite lidarmos com poucos parâmetros e poucos dados. Em compensação, o padrão de substituição é pouco razoável e com baixo grau de flexibilidade.

3 MÉTODOS E AJUSTES PARA A COMPARAÇÃO DOS MODELOS

3.1 Os Testes de Hipóteses

Quando temos um conjunto de modelos cujos os ajustes pretendemos comparar, o primeiro passo é identificar entre os pares de formas funcionais se há modelos aninhados, ou seja, se é possível chegar de um modelo a outro estabelecendo determinados valores para os coeficientes (Greene, 2003).

Dentre o modelos que estão sendo abordados no trabalho, temos duas relações de aninhamento. A primeira é entre o modelo PCAIDS e o modelo LAIDS. A partir do pressuposto de proporcionalidade, podemos do LAIDS chegar ao PCAIDS impondo as seguintes restrições:

$$\gamma_{jj} - \left(\frac{s_j}{1 - s_k} \frac{1 - s_j}{s_k} \gamma_{kk} \right) = 0 \quad (3.1)$$

$$\gamma_{jk} - \left(\frac{-s_j \gamma_{kk}}{(1 - s_k)} \right) = 0 \quad (3.2)$$

$$\beta_j = 0 \quad (3.3)$$

O modelo PCAIDS é considerado então aninhado ao modelo LAIDS.

A outra relação de aninhamento é entre o modelo LAIDS e log-linear. Quando restringimos os parâmetros $\beta_j = 0$, $\gamma_{jk} = 0$ e $\gamma = -1$ no modelo LAIDS, chegamos a um caso único do modelo log-linear – sendo essa então a única relação aninhada entre esses dois modelos, o que não é o suficiente para que um modelo seja considerado aninhado ao outro.

$$\eta_{jj} = [1/s_j(\gamma_{jj} - \beta_j s_k) + (1 + \beta_j/s_j)(1 + \gamma)s_j] - 1 \quad (3.4)$$

$$\eta_{jj} = -1 \quad (3.5)$$

Quando os modelos forem aninhados, é possível selecionar a forma funcional mais apropriada para descrever os dados a partir de um critério como F-test. Se não for possível derivar um modelo do outro através da imposição de restrições sobre parâmetros, teremos modelos não aninhados e será necessário utilizar testes estatísticos específicos.

Segundo Davidson e MacKinnon (2004), um método para comparar modelos não aninhados com a mesma variável dependente é o teste J. Proposto por esses mesmos autores em um artigo de 1981, ele é baseado na ideia da construção de um "modelo artificial", "ou aninhamento artificial". O teste será aplicado então da seguinte forma: suponha dois modelos, a hipótese nula (H_0) e a hipótese alternativa (H_1), que tentam explicar um mesma variável a partir de conjuntos de variáveis diferentes.

$$H_0 : y = f(X, \mu) + u_0 \quad (3.6)$$

$$H_1 : y = g(Z, \gamma) + u_1 \quad (3.7)$$

onde y é um n-vetor com o elemento típico y , sendo i a i -ésima observação, e as matrizes regressoras X e Z com proporção de n-linhas por k_1 -colunas e n-linhas por

k2-colunas, respetivamente, e u é um vetor de tamanho n relativo ao termo de erro e que $u_0 = NID(0, \sigma_0^2 I_T)$ e $u_1 = NID(0, \sigma_1^2 I_T)$.

Sendo λ um parâmetro escalar, podemos construir um modelo artificial da seguinte forma:

$$y = (1 - \lambda)f(X, \mu) + \lambda g(Z, \gamma) + u \quad (3.8)$$

Contudo, não é possível testar λ dessa forma pois há parâmetros que não podem ser identificados como afirmam Davidson e MAckinnon:

[... porque alguns dos regressores, ou combinação linear deles, podem aparecer em ambas as funções de regressão (Davidson e MAckinnon, 2004)]

Os autores propõem então a utilização de um valor previsto de g , \hat{g} , no lugar de $g(Z, \gamma)$ para contornar essa limitação. Chega-se, enfim, ao modelo artificial do teste J:

$$y = (1 - \lambda)f(X, \mu) + \lambda \hat{g} + u \quad (3.9)$$

O teste t pode ser usado para testar se $\lambda = 0$ é igual a zero. Se a hipótese $\lambda = 0$ não for rejeitada, podemos aceitar H_0 , ou seja, as variáveis não inclusas no H_0 não possuiriam poder explicativo a acrescentar (Greene 2003). As hipóteses são então invertidas, construímos um novo modelo artificial e realizamos um novo teste t sobre o λ . Como os testes são realizados de forma independente, o teste J pode

apresentar resultados inconclusivos, como todos os modelos sendo rejeitados ou todos os modelos sendo aceitos. Poderemos encontrar então os seguintes resultados para o teste:

	H0: $y=f(X, \mu)$	
H1: $y= g(Z, \gamma)$	Não rejeita	Rejeita
Não rejeita	Aceita f e g	Aceita g , rejeita f
Rejeita	Aceita f , rejeita g	Rejeita f e g

Quando uma das funções, $f(X, \beta)$ ou $g(Z, \gamma)$, for não linear, Davidson e Mackinnon sugerem a aplicação de uma alteração do teste J, o teste P (Davidson e Mackinnon, 1981). Para chegar ao teste P, começamos reescrevendo o modelo artificial do teste J, equação (3.9), da seguinte forma:

$$y = (1 - \lambda)f(X, \mu) + \lambda\hat{g} + u_0 \quad (3.10)$$

$$= f(X, \mu) + \lambda(\hat{g} - f) + u_0 \quad (3.11)$$

Nesse modelo é necessário estimar λ e μ ao mesmo tempo. O teste P é a linearização desse modelo artificial avaliado em $\mu = \hat{\mu}$ e $\alpha = 0$ com a seguinte equação para estimação

$$y - \hat{f} = \lambda(\hat{g} - \hat{f}) + \hat{F}b + u \quad (3.12)$$

onde \hat{F} é o vetor linha contendo as derivadas de f em relação a μ e $b = \mu - \hat{\mu}$. Nos casos em que a hipótese nula for linear, o teste J e teste P terão o mesmo valor (Davidson e Mackinnon, 1982). Isso porque

$$\hat{F}\hat{\mu} = \hat{f} \quad (3.13)$$

$$\hat{F}\mu = X\mu = f(X, \mu) \quad (3.14)$$

Dessa forma, poderemos chegar ao modelo artificial do teste J realizando as seguintes substituições:

$$y - \hat{f} = \lambda(\hat{g} - \hat{f}) + \hat{F}\mu - \hat{F}\hat{\mu} + u \quad (3.15)$$

$$y = X\mu + \lambda(\hat{g} - \hat{f}) + u \quad (3.16)$$

Se a variável dependente for diferente entre as equações, os testes P e J não poderão ser empregados. Davidson, White e Mackinnon (1983) sugerem que, caso as variáveis dependentes dos modelos não aninhados estejam submetidas à transformações como $\log y$, e^y , y^2 ou xy , onde x é uma variável exógena, utilize-se o teste PE, uma extensão do teste P (Davidson e Mackinnon, 1983). O teste PE pode ser aplicado da seguinte forma: sendo $h(\cdot)$ uma função monótona, continuamente diferenciável e não dependente de parâmetros desconhecidos, suponha os modelos:

$$H_0 : y = f(X, \mu) + u_0 \quad (3.17)$$

$$H_1 : h(y) = g(Z, \gamma) + u_1 \quad (3.18)$$

Assim como fizemos para o teste J, podemos construir o seguinte modelo artificial:

$$y - \hat{f} = \lambda(\hat{g} - h(y)) + \hat{F}b + u \quad (3.19)$$

Uma regressão direta teria termos envolvendo y no lado direito. Andrews (1971), apud Mackinnon, White and Davidson (1983), afirma que é possível substituir y por \hat{f} toda vez que ele aparecer no lado direito. Seguindo esse procedimento temos:

$$y - \hat{f} = \lambda(\hat{g} - h(\hat{f})) + \hat{F}b + u \quad (3.20)$$

Assim como ocorre para o teste P, equação (3.16), se as hipóteses forem lineares podemos utilizar uma forma mais simplificada do modelo para realizar o teste PE:

$$y = X\mu + \lambda(\hat{g} - h(\hat{f})) + u_i \quad (3.21)$$

Podemos observar que se estivermos lidando com modelos lineares, como é o caso dos modelos que estão sendo estudados neste trabalho, e a transformação da equação (3.21) for igual a um, os modelos compostos dos testes J, P e PE serão iguais e, portanto, os resultados dos testes serão os mesmos. Dessa forma, doravante, iremos nos referir aos três testes apenas como teste PE, já que os demais são apenas casos especiais do último.

Outro teste que pode também ser aplicado aos modelos com variável dependente transformada é o teste BEM. Proposto por Bera and McAleer (1989), esse teste segue a mesma abordagem de "modelo artificial". A partir das hipóteses mencionadas nas equações (3.17) e (3.18), o autores propõem que ao invés de ser usado o valor previsto da variável dependente com a qual é construído o modelo artificial,

seja utilizado o valor previsto dos resíduos. Além disso, ao invés de utilizarem a transformação sobre a variável prevista do próprio modelo, como no caso do termo $h(\hat{f})$ do teste PE (equação(3.21)), eles transformam a hipótese alternativa \hat{g} para que o resíduo dela fique na mesma escala do modelo no qual ele será inserido. Para realizar o teste BEM são necessários três passos. Primeiro é necessário estimar os sistemas (3.17), (3.18) e obter assim os valores previstos \hat{y} e $h(\hat{y})$

$$\hat{y} = f(X, \mu) + \mu_0 \quad (3.22)$$

$$h(\hat{y}) = g(Z, \gamma) + \mu_1 \quad (3.23)$$

No segundo passo, aplicamos a transformação $h()$ e sua inversa sobre os modelos de maneira respectiva e chegamos nos modelo que iremos estimar para obter η_0 e η_1 abaixo

$$h(\hat{y}) = g(Z, \gamma) + \eta_1 \quad (3.24)$$

$$h^{-1}[h(\hat{y})] = f(X, \mu) + \eta_0 \quad (3.25)$$

Por fim, inserimos os valores obtidos $\hat{\eta}_0$ e $\hat{\eta}_1$ em (3.18) e (3.17), respectivamente, e chegamos aos modelos que serão estimados e sobre os quais realizaremos um teste de t para testar $\Theta_0 = 0$ e $\Theta_1 = 0$, ou seja, testar se o modelo adicionado é relevante para explicar os dados.

$$y = f(X, \mu) + \Theta_0 \eta_1 + \epsilon \quad (3.26)$$

$$h(y) = g(Z, \gamma) + \Theta_1 \eta_0 + \epsilon \quad (3.27)$$

A motivação para usarmos também o teste BEM, além do teste PE, é o artigo Manera (1991). Nesse trabalho, o autor realizou uma aplicação empírica e uma análise de Monte Carlo para verificar o desempenho de ambos os testes. Os resultados foram mais favoráveis ao teste BEM, apesar do autor afirmar que os dois testes seriam úteis .

Manera (1991) compara cinco sistemas de demanda de fatores não aninhados: *Translog Factor Demand System* (TLG), *Quadratic Factor Demand System* (QDR), *Generalized Leontief Factor Demand System* (GLT), *Symmetric Generalized McFadden Factor Demand System* (SGM), *Flexible Functional Forms Factor Demand System* (FFF) e *Quadratic Non-Homogeneous Factor Demand System* (NHQDR) - sendo apenas dois deles com a mesma variável dependente. Ao aplicar empiricamente os testes, tanto o PE e quanto o BEM chegam aos mesmos resultados em uma comparação pareada. O modelo TLG seria preferível ao QDR, enquanto que NHQDR se sairia pior que SGM e o GLT.

Contudo, na análise de Monte Carlo realizada por Manera(1991), o teste PE não apresentou resultados satisfatórios, em particular quando a hipótese nula é verdadeira, apresentando erros do tipo 1 muito frequentemente. Os testes foram realizados comparando os modelos QDR e TLG e, em uma amostra do tamanho de 25, o PE-teste rejeitou metade das vezes a hipótese nula quando ela era correta. A situação piora no caso de modelos como SGM e NHQDR. Nesse caso, o teste PE rejeita aproximadamente 65% das hipóteses nulas verdadeiras. No caso de uma amostra com tamanho de 100, o teste PE rejeita a hipótese verdadeira muitas vezes, 42 e 58 por cento das vezes para os respectivos modelos sendo comparados. O teste se sai melhor em relação ao erro do tipo 2, ou seja, para a probabilidade em rejeitar uma hipótese quando ela é falsa. O teste PE para os modelos GLT e NHQDR, numa amostra de tamanho 25, rejeitou a hipótese nula quando ela era falsa aproximadamente 78 por cento das vezes. Quando a amostra aumenta para

100, os resultados melhoram significativamente e o número de vezes que o PE rejeita a hipótese nula é igual ao tamanho nominal do teste.

Todos os testes mostrados até o momento foram explicados em sua forma univariada, de acordo como foram apresentados nos artigos originais. Todavia, sistemas de demanda são modelos multivariados. A não ser que as equações de demanda sejam estimadas na forma pooled, num estudo com j produtos teremos j modelos artificiais para serem testados. Contudo, como mostra Manera (1991,2002,2005), esses mesmos testes podem ser estendidos para a forma multivariada de maneira simples. Ele indica a utilização de um teste de Wald para testar se $\lambda_j = \dots = \lambda_j = 0$ da equação (3.21), ou $\Theta_0 1 = \dots = \Theta_0 j = 0$ na equação (3.26) e $\Theta_1 1 = \dots = \Theta_1 j = 0$ de (3.27), são iguais a zero ao mesmo tempo.

3.2 A implementação dos Testes para Modelos de Demanda

3.2.1 Comparação do Modelo Log-Linear e Logit

Para comparar o modelo log-linear e Logit é conveniente igualar as variáveis dependentes. O modelo Logit pode ser escrito em função da quantidade da seguinte forma:

$$\ln(w_j) - \ln(w_0) = \ln(Q_j) - \ln(Q_0) = -\pi P_j + D_j \omega + \xi_j + \epsilon_j \quad (3.28)$$

$$\ln(Q_j) = -\pi P_j + D_j \omega + \xi_j + \ln(Q_0) + \epsilon_j \quad (3.29)$$

Apesar dos dois modelos apresentarem $\ln(Q_j)$ como variável dependente, o lado esquerdo da equação do modelo Logit apresenta também $\ln(Q_0)$. Uma saída

então é passar $\ln(Q_0)$ para o lado esquerdo da equação e considerá-lo uma variável explicativa. Para isso, será necessário impor a restrição de que o coeficiente de variável $\ln(Q_j)$ seja igual a 1. O modelo artificial do teste PE, com o modelo log-linear como hipótese nula e o modelo Logit como hipótese alternativa, foi construído a partir das equações (2.2), (3.21) e (3.29), seguindo a seguinte forma:

$$\log Q_j = \alpha_{j0} + \sum_k \alpha_{jk} \log P_k + \alpha_{jD} \log D + \lambda_j \left(\ln Q_{j(\widehat{Logit})} - \ln Q_{j(\widehat{log-linear})} \right) + \epsilon_j \quad (3.30)$$

E ao invertermos as hipóteses temos

$$\ln Q_j = -\pi P_j + D_j \omega + \xi_j + \ln(Q_0) + \lambda_j \left(\ln Q_{j(\widehat{Logit})} - \ln Q_{j(\widehat{log-linear})} \right) + \epsilon_j \quad (3.31)$$

3.2.2 Comparação do Modelo Log-linear e LAIDS ou PCAIDS

No caso do modelo log-linear e LAIDS, temos um com o log da quantidade do produto j e o outro com o share de valor. Para compararmos os modelos através do teste PE é necessário aplicar uma das seguintes transformações:

$$h(\ln Q_j) = \frac{(e^{\ln Q_j} * P_j)}{Y} \quad (3.32)$$

$$h(s_j) = \ln \frac{(s_j * Y)}{P_j} \quad (3.33)$$

Para realizar o teste PE, tendo como hipótese nula o modelo log-linear e como hipótese alternativa o modelo LAIDS, foi montado um modelo artificial para cada um dos três produtos utilizando as equações (2.2), (2.6), (3.21) e (3.32). Chegamos então na seguinte equação para ser estimada:

$$\log Q_j = \alpha_{j0} + \sum_k \alpha_{jk} \log P_k + \alpha_{jD} \log D + \lambda_j \left(\hat{s}_j - \frac{(e^{\ln \hat{Q}_j} * P_j)}{Y} \right) + \epsilon_j \quad (3.34)$$

Ao invertermos as hipóteses, utilizou-se as equações (2.6), (2.2), (3.21) e (3.33).

$$s_j = \beta_{0j} + \beta_j \log \left(\frac{Y}{I} \right) + \sum_k \gamma_{jk} \log P_k + D\Theta_j + \lambda_j \left(\ln \hat{Q}_j - \ln \frac{(\hat{s}_j * Y)}{P_j} \right) + \epsilon_j \quad (3.35)$$

Para o teste BEM, foram utilizadas as equações (2.2), (2.6), (3.26),(3.27), (3.32) e (3.33) para construir os modelos para estimar o segundo passo do teste BEM

$$h(\log \hat{Q}_j) = \beta_{0j} + \beta_j \log \left(\frac{Y}{I} \right) + \sum_k \gamma_{jk} \log P_k + D\Theta_j + \eta_{0j} \quad (3.36)$$

$$h^{-1}(\hat{s}_j) = \alpha_{j0} + \sum_k \alpha_{jk} \log P_k + \alpha_{jD} \log D + \eta_{1j} \quad (3.37)$$

E o último passo

$$\log Q_j = \alpha_{j0} + \sum_k \alpha_{jk} \log P_k + \alpha_{jD} \log D + \Theta_{0j} \hat{\eta}_{1j} + \epsilon_j \quad (3.38)$$

$$s_j = \beta_{0j} + \beta_j \log \left(\frac{Y}{I} \right) + \sum_k \gamma_{jk} \log P_k + D\Theta_j + \Theta_{1j} \hat{\eta}_{0j} + \epsilon_j \quad (3.39)$$

3.2.3 Comparação Modelo Logit e LAIDS ou PCAIDS

Da mesma forma que na comparação com o modelo log-linear, será necessário rescrever as variáveis dependentes do modelo Logit em função das quantidades e passar a relacionada ao produto externo para o lado esquerdo e considerá-la igual a 1. O modelo LAIDS terá como variável dependente o logaritmo em função do share de valor. A partir daí, será necessária uma das transformações já apresentadas, (3.32) e (3.33). No caso do Modelo Logit e LAIDS, considerando como hipótese nula e alternativa respectivamente os modelos citados, chegou-se ao modelo artificial através das equações (3.29), (3.21) e (3.32) para cada um dos produtos.

$$\ln(Q_j) = -\pi P_j + D_j \omega + \xi_j + \ln(Q_0) + \lambda_j \left(\hat{s}_j - \frac{(e^{\ln \hat{Q}_j} * P_j)}{Y} \right) + \epsilon_j \quad (3.40)$$

E as equações (2.6), (3.21) e (3.33) ao considerarmos como hipótese nula o modelo LAIDS e hipótese alternativa o modelo Logit.

$$s_j = \beta_{0j} + \beta_j \log \left(\frac{Y}{I} \right) + \sum_k \gamma_{jk} \log P_k + D\Theta_j + \lambda_j \left(\ln \hat{Q}_j - \ln \frac{(\hat{s}_j * Y)}{P_j} \right) + \epsilon_j \quad (3.41)$$

Para o teste BEM, foram utilizadas as equações (3.29), (2.6), (3.26), (3.27), (3.32) e (3.33) para o primeiro passo ao estimar

$$h(\log \hat{Q}_j) = \beta_{0j} + \beta_j \log \left(\frac{Y}{I} \right) + \sum_k \gamma_{jk} \log P_k + D\Theta_j + \eta_{0j} \quad (3.42)$$

$$h^{-1}(\hat{s}_j) = -\pi P_j + D_j \omega + \xi_j + \ln(Q_0) + \eta_{1j} \quad (3.43)$$

E último passo

$$\log Q_j = -\pi P_j + x_j \omega + \xi_j + \ln(Q_0) + \Theta_{0j} \hat{\eta}_{1j} + \epsilon_j \quad (3.44)$$

$$s_j = \beta_{0j} + \beta_j \log \left(\frac{Y}{I} \right) + \sum_k \gamma_{jk} \log P_k + D\Theta_j + \Theta_{1j} \hat{\eta}_{0j} + \epsilon_j \quad (3.45)$$

A tabela 3.1 apresenta, de maneira simplificada, os métodos que adotaremos e os ajustes necessários entre os modelos:

Tabela 3.1: Comparação de modelos

Hipótese Alternativa (H1)	Hipótese Nula (H0)			Logit
	Log-linear	AIDS	PCAIDS	
Log-linear		Não-aninhado Caso especial: $\beta_j = 0, \gamma_{jk} = 0$ e $\gamma = -1$ Teste: PE-test Transformação: $h(s_j) = \ln \frac{(s_j * Y)}{p_j}$	Não-aninhado Teste: PE-test Transformação: $h(s_j) = \ln \frac{(s_j * Y)}{p_j}$	Não-aninhado Teste: J-test
LAIDS	Não-aninhado Teste: PE-test Transformação: $h(lnq) = \frac{(e^{ln s_j * p_j})}{Y}$	Transformação: $h(s_j) = \ln \frac{(s_j * Y)}{p_j}$	Aninhados Teste F	Não-aninhado Teste: PE-test Transformação: $h(lnq) = \frac{(e^{ln s_j * p_j})}{Y}$
PCAIDS	Não-aninhado Teste: PE-test Transformação: $h(lnq) = \frac{(e^{ln s_j * p_j})}{Y}$	Aninhados Teste F-test Restrições: $\beta_j = 0$ $\gamma_{jj} = \left(\frac{s_j}{1-s_k} \frac{1-s_k}{s_k} \gamma_{kk} \right)$ $\gamma_{jk} = \left(\frac{s_j}{(1-s_k) p_{jk}} \right)$		Não-aninhado Teste: PE-test Transformação: $h(lnq) = \frac{(e^{ln s_j * p_j})}{Y}$
Logit	Não-aninhados Teste: J-test Imposição: $\ln q_0$ como variável explicativa e igual a 1.	Não-aninhado Teste: PE-test Transformação: $h(s_j) = \ln \frac{(s_j * Y)}{p_j}$	Não-aninhado Teste: PE-test Transformação: $h(s_j) = \ln \frac{(s_j * Y)}{p_j}$	

4 DADOS

Para demonstrar a aplicabilidade da metodologia proposta e comparar a capacidade de explicação da curva de demanda dos modelos apresentados no capítulo 2, foi realizada uma aplicação com os dados semanais de produtos diferenciados da Dominick's Finer Foods, uma grande rede varejista dos EUA. Os dados estão disponíveis publicamente no site da University of Chicago Booth School of Business, pelo James M. Kilts Center. A base de dados é um painel constituído de mais de 3.500 produtos e conta com informações como nome do produto, tamanho, preço, taxa de lucro e quantidade vendida por semana para cada um, em cada loja, por um período de cerca de 390 semanas.

Para o trabalho, foi selecionado como produto agregado “refrigerantes de cola de dois litros normais”, considerando como “normais” os refrigerantes que não sejam dietéticos ou com algum tipo de sabor adicional, e selecionadas as três marcas com maior quantidade vendida no período total da amostra para representar os bens internos. As marcas encontradas, em ordem decrescente de quantidade vendida no período total, foram : Pepsi (produto 1); Coca – Cola (produto 2); Domincks Cola Soda (produto 3).

Consideramos como produto externo, necessário para estimação do modelo Logit, todos os refrigerantes da base menos os produtos selecionados para bens internos. A exemplo de DeSouza (2009), o preço do bem externo é assumido como exógeno e não interage com as decisões de preços dos bens internos.

O market share em função do valor(s_j) foi obtido pela seguinte equação

$$s_j = \frac{P_j Q_j}{\sum_{k=1}^N P_k Q_k} \quad (4.1)$$

E o market share em função da quantidade (w_j) por

$$w_j = \frac{Q_j}{\sum_{k=0}^N Q_k} \quad (4.2)$$

A variável de custo de cada produto, (c_j), foi obtida através da multiplicação do preço médio de venda de cada produto (P_j) por (1-profit) ,sendo “profit” a variável de lucro marginal em porcentagem disponibilizada pela base.

O deslocador de demanda D foi considerado como sendo o dispêndio nos três produtos no período, Y . A função logarítmica foi aplicada sobre os preços ($\ln P_j$), quantidades ($\ln Q_j$) e renda ($\ln Y$). Os valores das variáveis estão sumarizados na tabela 4.1.

Tabela 4.1: Estatísticas das variáveis utilizadas

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
M	390	13,772	11,209	3,223	121,240
Y	390	728.9	551.3	117.5	3,674
c1	390	1.200	0.289	0.176	2.841
c2	389	1.136	0.253	0.178	2.670
c3	390	0.579	0.265	0.173	1.745
p1	390	1.396	0.299	0.790	2.900
p2	389	1.376	0.288	0.772	2.888
p3	390	0.915	0.320	0.390	2.133
q0	390	13,116	11,159	2,002	120,673
q1	390	319.9	410.0	51.62	3,346
q2	389	264.5	295.3	54.76	1,910
q3	390	72.80	108.0	6.662	741.2
s1	390	0.478	0.160	0.0508	0.952
s2	389	0.424	0.156	0.0439	0.944
s3	390	0.0995	0.106	0.00227	0.466
w1	390	0.0339	0.0542	0.000653	0.480
w2	389	0.0308	0.0496	0.000742	0.394
w3	390	0.00541	0.00493	0.000129	0.0262
week	391	198.8	116.2	1	399
lnY	390	6.388	0.602	4.766	8.209
lnc1	390	0.151	0.269	-1.735	1.044
lnc2	389	0.0960	0.276	-1.726	0.982
lnc3	390	-0.612	0.331	-1.754	0.557
lnp1	390	0.311	0.212	-0.236	1.065
lnp2	389	0.298	0.211	-0.259	1.060
lnp3	390	-0.142	0.323	-0.942	0.757
lnq0	390	9.232	0.702	7.602	11.70
lnq1	390	5.279	0.902	3.944	8.116
lnq2	389	5.168	0.829	4.003	7.555
lnq3	390	3.682	0.976	1.896	6.608

Fonte: Autoria própria.

5 APLICAÇÃO EMPÍRICA

5.1 Métodos de Estimação dos Modelos

Os sistemas de demanda são considerados modelos de regressão lineares multivariados, onde temos N equações com mesmas variáveis dependentes e mesmas variáveis explicativas entre elas. Quando os termos de erro das equações do modelo forem independentes, como no caso do modelo log-linear, cada equação do sistema poderá ser estimada de maneira eficiente por Mínimos Quadrados Ordinários (MQO) (Davidson e Mackinnon, 1999). Contudo, em casos como do modelo LAIDS, em que as variáveis dependentes são as parcelas de mercado s_j e que o somatório dessas deve que ser igual a um, e do Logit, que também impõe restrições entre os parâmetros, os erros das equações não são independentes entre si. Para esses dois modelos, a estimação por mínimos quadrados produzirá coeficientes ineficientes. O método mais adequado e de fácil implementação então é o *Seemingly Unrelated Regressions* (SUR), apresentado por Zellner (1962), que faz uso das correlações dos erro entre equações (Greene, 2003).

Um tipo de má especificação que ocorre em modelos de demanda, e que será tratada no trabalho, é a determinação conjunta de preço e quantidade em um mercado. A presença de simultaneidade corrompe uma das premissas do MQO, e, assim, do SUR, ocasionando correlação entre os preços e o erro da regressão. A estimação por um desses métodos gera coeficientes relacionados aos preços inconsistentes e as elasticidades, calculadas a partir deles, subestimadas em valores absolutos (DEE, 2009).

O problema da endogeneidade da variável preço pode ser corrigido através do uso de uma variável instrumental. Para uma variável ser considerado um instrumento ela deve estar correlacionada à variável explicativa endógena e não estar correlacionada ao erro. No caso de curvas de demanda, utilizam-se deslocadores de oferta como instrumentos. A base de dados escolhida para o trabalho permite a construção de uma variável de custo, como mostrado na seção 5, supostamente exógena para o supermercado dado o tamanho do varejo nos EUA, e que será usada para esse fim. Podemos verificar a existência de correlação entre o erro e alguma das variáveis explicativas, ou seja, se há presença de endogeneidade, através do teste de endogeneidade de Hausman e Durbin. Considerando a variável explicativa relacionada ao preço sendo exógena como hipótese nula, as estatísticas rejeitaram essa possibilidade em todos os modelos (Apêndice A). Outro teste que pode ser realizado é o para verificar se os instrumentos são fracos. Nesse caso, essa hipótese foi rejeitada, como pode ser visto no Apêndice B. Os métodos correspondentes ao MQO e ao SUR que empregam variáveis instrumentais são, respectivamente, o método de Mínimos Quadrados em Dois Estágios (MQ2E) e o método de Mínimos Quadrados em Três Estágios (MQ3E) ou Método Geral dos Momentos Generalizado (MMG) - que acomoda o MQ3E como caso particular DEE(2009).

Assim como a estimação dos modelos de demanda, a estimação dos modelos artificiais para os testes que serão implementados nesse trabalho podem sofrer também de inconsistência se o método de estimação escolhido não corrigir uma possível correlação entre as variáveis de preço e o erro ϵ_j . Godfrey (1983) afirma que os testes propostos por Davidson e Mackinnon só seriam válidos após a estimação por variável instrumental se o modelo da hipótese alternativa tivesse apenas uma variável explicativa e que não fosse explicada pela combinação linear dos repressores do outro modelo. Contudo, Davidson, White e Mackinnon (1983) explicam que o teste J e o teste P podem ser modificados para lidar com essa situação e demonstram como empregá-lo. Os autores, por simplicidade, se detêm a um caso de modelo linear e

univariado.

Utilizando os modelos vistos anteriormente em (3.6) e (3.7), considere que algumas das colunas de X e Z podem ser correlacionadas com μ_0 e μ_1 e que exista uma matriz de instrumentos, W , com as propriedades necessárias a um instrumento, de forma que o MQ2E é possível. Os autores assumem também que os dois modelos competitivos possuem a mesma matriz de instrumentos - o que é compatível com o nosso caso pois as explicativas endógenas são as mesmas, ou seja, preços. Sendo P_W a projeção de W , a estimação de mínimos quadrados em dois estágios de H_0 e H_1 pode ser obtida através da estimação por MQO

$$y = P_W X \beta + \mu_0 \quad (5.1)$$

e

$$y = P_W Z \gamma + \mu_1 \quad (5.2)$$

Onde $P_W = W(W^T W)^{-1} W^T$. Os coeficientes de MQ2E serão

$$\hat{\beta} = (X^T P_W X)^{-1} X^T P_W y \quad (5.3)$$

$$\hat{\gamma} = (Z^T P_W Z)^{-1} Z^T P_W y \quad (5.4)$$

O seguinte modelo artificial para o calculo da estatística, J no caso, poderá ser estimado por MQO

$$y = P_W X \beta + P_{ZW} y + \mu, \quad (5.5)$$

em que $P_{ZW} = P_W Z (Z^T P_W Z)^{-1} Z^T P_W$.

A estrutura para aplicação da variável instrumental pode ser seguida para casos multivariados e para o teste PE, que é apenas uma extensão do teste P. Não foi encontrado na literatura alguma menção sobre o emprego de variáveis instrumentais no caso do teste BEM. Contudo, como se trata de uma teste que segue o mesmo princípio de aninhamento artificial, não parece haver problema em utilizar as variáveis instrumentais da mesma forma que nos testes do Davidson e Mackinnon.

5.2 Estimação dos modelos

Para demonstrar a aplicabilidade da metodologia proposta e indicar qual dos modelos apresentados no capítulo 2 explicaria melhor a curva de demanda, estimaremos os modelos com o dados do capítulo 4 para depois computarmos as estatísticas.

5.2.1 Modelo Log-linear

Nas tabelas 5.1 e 5.2 apresentamos os resultados para a estimação do modelo log-linear pelo método SUR e MQ3E, respectivamente. Os coeficientes encontrados para as estimações sem o auxílio de variável instrumental foram todos significativos ao nível de 1%, com exceção ao coeficiente referente a equação 2 com o preço do produto 3, γ_{23} (-0,0181), que é não significativo.

Tabela 5.1: Modelo log-linear estimado por SUR

	(1)	(2)	(3)
VARIABLES	lnq1	lnq2	lnq3
lnp1	-2.491*** (0.105)	1.088*** (0.1000)	0.736*** (0.260)
lnp2	1.339*** (0.104)	-2.649*** (0.0992)	1.032*** (0.258)
lnp3	0.151*** (0.0522)	-0.0181 (0.0498)	-1.373*** (0.129)
lnY	0.994*** (0.0360)	0.767*** (0.0343)	0.143 (0.0892)
Constant	-0.675*** (0.254)	0.714*** (0.242)	2.038*** (0.630)
Observations	389	389	389
R-squared	0.865	0.854	0.289

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Fonte: Autoria própria.

A estimação com o uso de variável instrumental apresentou um número maior de coeficientes não significativos: γ_{21} (-0,0243), γ_{13} (0,660), γ_{32} (1,252) . Os demais são significativos ao menos pelo nível de significância de 5%.

Tabela 5.2: Modelo log-linear estimado por MQ3E

VARIABLES	(1) lnq1	(2) lnq2	(3) lnq3
lnp1	-1.514*** (0.264)	-0.0243 (0.268)	0.660 (0.729)
lnp2	0.617** (0.291)	-1.796*** (0.296)	1.252 (0.805)
lnp3	-0.00878 (0.0843)	-0.250*** (0.0856)	0.566** (0.233)
lnY	1.052*** (0.0514)	0.721*** (0.0523)	0.104 (0.142)
Constant	-1.159*** (0.382)	1.067*** (0.388)	2.522** (1.058)
Observations	389	389	389
R-squared	0.828	0.790	-0.124

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1
 Fonte: Autoria própria.

5.2.2 Modelo LAIDS

O modelo LAIDS foi estimado através da imposição das restrições de simetria e homogeneidade dos parâmetros entre as equações. Os valores dos parâmetros da equação do quarto produto foram obtidos através da combinação linear pelas seguintes equações:

$$\sum_{j=1}^4 \beta_{0j} = 1 \quad (5.6)$$

$$\sum_{j=1}^4 \beta_j = 0 \quad (5.7)$$

$$\sum_{j=1}^4 \gamma_{4j} = 0 \quad (5.8)$$

Os coeficientes provenientes das regressões do modelo estimados pelo método SUR, tabela 5.3, são todos significativos a um nível de significância de 5%, com exceção de apenas três deles: γ_{23} (0,0196), γ_{33} (0,0196) e γ_{β_2} (-0,00324) .

Tabela 5.3: Modelo LAIDS estimado por SUR

VARIABLES	s1	s2	s3
lnp1	-0.562*** (0.0318)	0.516*** (0.0283)	0.0460*** (0.0152)
lnp2	0.516*** (0.0283)	-0.535*** (0.0307)	0.0196 (0.0148)
lnp3	0.0460*** (0.0152)	0.0196 (0.0148)	-1.123*** (0.0635)
lnindex	0.0308*** (0.00859)	-0.00324 (0.00848)	-0.027*** (0.0073)
Constant	0.315*** (0.0557)	0.445*** (0.0548)	0.239*** (0.0478)
Observations	389	389	389
R-squared	0.446	0.441	-

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Fonte: Autoria própria.

Quando o mesmo modelo é estimado através do MQ3E, 5.4, os coeficientes são todos significativos a 5%, com exceção de γ_{13} (-0,0164), γ_{31} (-0,0164), γ_{β_2} (0,0172).

Tabela 5.4: Modelo LAIDS estimado por MQ3E

VARIABLES	s1	s2	s3
lnp1	-0.321*** (0.0848)	0.337*** (0.0853)	-0.0164 (0.0274)
lnp2	0.337*** (0.0853)	-0.229** (0.0930)	-0.108*** (0.0299)
lnp3	-0.0164 (0.0274)	-0.108*** (0.0299)	-0.641*** (0.1697)
lnindex	0.0413*** (0.00988)	0.0172 (0.0105)	-0.058*** (0.0094)
Constant	0.219*** (0.0676)	0.266*** (0.0720)	0.515*** (0.0654)
Observations	389	389	389
R-squared	0.368	0.264	-

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Fonte: Autoria própria.

5.2.3 Modelo PCAIDS

Os coeficientes do modelo PCAIDS foram encontrados utilizando a proposta do Coloma(2006). Para simular um caso em que poderíamos aplicar o modelo, o autor supõe um cenário onde temos informações sobre o preço de apenas de dois produtos, P1 e P2, e os market shares de todos produtos. O sistema que antes era escrito a partir de todos os preços, utilizando as equações (2.6), deve passar a depender apenas de P1 e P2. Para isso, será necessário inverter o sistema deixando os preços como variáveis dependentes, o que necessita que a matriz formada pelos coeficientes γ_{ij} , Matriz A, seja invertida originando uma nova matriz, Matriz B.

$$A = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{bmatrix} \quad (5.9)$$

$$= \begin{bmatrix} \gamma_{11} & \frac{-s_2}{(1-s_1)}\gamma_{11} & \frac{-s_3}{(1-s_1)}\gamma_{11} \\ \frac{-s_2}{(1-s_1)}\gamma_{11} & \frac{s_2}{1-s_1} \frac{1-s_2}{s_1} \gamma_{11} & \frac{-s_2 s_3}{s_1(1-s_1)}\gamma_{11} \\ \frac{-s_3}{(1-s_1)}\gamma_{11} & \frac{-s_2 s_3}{s_1(1-s_1)}\gamma_{11} & \frac{s_3}{1-s_1} \frac{1-s_3}{s_1} \gamma_{11} \end{bmatrix} \quad (5.10)$$

Como a matriz A é uma matriz singular, é necessário eliminar uma coluna e uma linha, de algum dos produtos que temos informação, para podermos invertê-la. Eliminando as linhas e colunas ao produto 2, chegaremos à matrix B. Encontraremos então os novos coeficientes:

$$B^{-1} = \begin{bmatrix} \nu_{11} & \nu_{13} \\ \nu_{31} & \nu_{33} \end{bmatrix} \quad (5.11)$$

$$= \begin{bmatrix} \frac{(s_1+s_2)(1-s_1)}{\gamma_{11}s_2} & \frac{s_1(1-s_1)}{\gamma_{11}s_2} \\ \frac{\gamma_{11}s_2}{s_1(1-s_1)} & \frac{(s_3+s_2)s_1(1-s_1)}{\gamma_{11}s_2s_3} \end{bmatrix} \quad (5.12)$$

Com as seguintes expressões gerais para os coeficientes:

$$\nu_{jj} = \frac{(s_j + s_2)s_1(1 - s_1)}{\gamma_{11}s_2s_j} \quad (5.13)$$

$$\nu_{jk} = \frac{s_1(1 - s_1)}{\gamma_{11}s_2} \quad (5.14)$$

A partir das equações (5.13), (5.14) e da premissa $\beta_j = 0$ do modelo PCAIDS, o sistema (2.9) pode ser reescrito então como:

$$\ln(P_j) = \nu_{j0} + \nu_{j1}s_1 + \nu_{j3}s_3 + \dots + \nu_{jj}s_j + \nu_{j2}\ln(P_2) + \nu_{jD}D \quad (5.15)$$

Como a linha e coluna relativas ao produto dois foram excluídas da matriz, ν_{j1} . Substituindo ν_{jj} e ν_{jk} temos

$$\ln(P_j) = \nu_{j0} + \frac{(s_j + s_2)s_1(1 - s_1)s_1}{\gamma_{11}s_2s_1} + \frac{s_1(1 - s_1)(1 - s_1 - s_2)}{\gamma_{11}s_2} + \nu_{j2}\ln(P_2) + \nu_{jD}D \quad (5.16)$$

$$\ln(P_j) = \nu_{j0} + \frac{s_j(1 - s_1)}{\gamma_{11}s_2} + \ln(P_2) + \nu_{jD}D \quad (5.17)$$

Essa equação pode ser aplicada a todos os preços, contudo, seguindo a suposição de Coloma de que temos apenas os preços dos produtos 1 e 2, teremos à disposição apenas uma equação para estimarmos γ_{11} :

$$\ln(p_1) = \nu_{10} + \frac{s_1(1 - s_1)}{\gamma_{11}s_2} + \ln(P_2) + \nu_{1D}D \quad (5.18)$$

Ou de forma para se obter uma estimativa direta:

$$\frac{s_1(1 - s_1)}{s_2} = -\gamma_{11}\nu_{10} + \gamma_{11}\frac{\ln(P_1)}{\ln(P_2)} - \gamma_{11}\nu_{1D}D \quad (5.19)$$

Tabela 5.5: Modelo PCAIDS estimado por MQO

VARIABLES	(1) MQO
g11	0.00526*** (0.00108)
Constant	0.605*** (0.0107)
Observations	389
R-squared	0.058

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Fonte: Autoria própria.

Para a aplicação empírica em um mercado com três grandes marcas, através das equações (5.19) e (2.6), chegamos a equação que estimaremos

$$\frac{s_1(1 - s_1)}{s_2} = \beta_0 + \gamma_{11} \frac{\ln(P_1)}{\ln(P_2)} + D\Theta + \epsilon \quad (5.20)$$

O coeficiente único encontrado a partir da estimação realizada por MQO, tabela 5.5, produziu um coeficiente com o sinal positivo (0,005) e estatisticamente significativo a 1%. Para lidar com problema da endogeneidade relacionada às variáveis de preço, foram usados como instrumentos os custos dos três produtos e a estimação realizada pelo método de Mínimo Quadrado em Dois Estágios (MQ2E). O coeficiente encontrado, contudo, não foi significativo, como mostra a tabela 5.6.

Após estimar γ_{11} , podemos chegar aos demais coeficientes substituindo os valores nas equações (2.10) e (2.11). Os coeficientes relativos aos preços obtidos utilizando os dados do capítulo 4 e estimação por MQO são apresentados na tabela 5.7 e por MQ2E na tabela 5.8.

Tabela 5.6: Modelo PCAIDS estimado por MQ2E

VARIABLES	(1) MQ2E
g11	-0.102 (0.138)
Constant	0.603*** (0.0552)
Observations	389

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1
 Fonte: Aatoria própria.

Tabela 5.7: Modelo PCAIDS - Coeficientes obtidos a partir da estimação de γ_{11} por MQO.

	g1	g2	g3
1	0.005*	-0.004*	-0.001*
2	-0.004*	0.005*	-0.001*
3	-0.001*	-0.001*	0.002*

Fonte: Aatoria própria.

Tabela 5.8: Modelo PCAIDS - Coeficientes obtidos a partir da estimação de γ_{11} por MQ2E.

	g1	g2	g3
1	-0.102	0.083	0.019
2	0.083	-0.100	0.017
3	0.019	0.017	-0.037

Fonte: Aatoria própria.

5.2.4 Modelo Logit

Para testarmos as equações de um único produto entre modelos, passamos o modelo Logit para a forma de um sistema com três equações com a imposição de que os coeficientes de preços sejam todos iguais - o que justifica aplicação do método SUR. Consideramos como característica dos produtos as marcas dos bens internos. Os resultados da estimação por SUR foi apresentados na tabela 5.9 e para a estimação por MQ3E na tabela 5.10 . Como mencionado, a variável relacionada à quantidade do bem externo é considerada um efeito fixo e, por se tratar de uma função logarítmica, fica igual a 1. Os coeficientes do modelo estimado sem e com variável instrumental apresentaram sinais negativos, -1,809 e -0,458, e significativos a 1%.

Tabela 5.9: Modelo Logit estimado por SUR

VARIABLES	(1) lnq1	(2) lnq2	(3) lnq3
p1	-1.809*** (0.0889)		
p2		-1.809*** (0.0889)	
p3			-1.809*** (0.0889)
Constant	-1.427*** (0.134)	-1.575*** (0.131)	-3.893*** (0.0969)
Observations	389	389	389
R-squared	-0.180	-0.244	-0.240

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1
 Fonte: Autoria própria.

Tabela 5.10: Modelo Logit estimado por MQ3E

VARIABLES	(1) lnq1	(2) lnq2	(3) lnq3
p1	-0.458*** (0.132)		
p2		-0.458*** (0.132)	
p3			-0.458*** (0.132)
Constant	-3.312*** (0.191)	-3.434*** (0.188)	-5.129*** (0.130)
Observations	389	389	389
R-squared	-0.456	-0.644	0.025

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1
 Fonte: Autoria própria.

5.3 Elasticidades

Para encontrar as elasticidades do modelo LAIDS e do PCAIDS, consideramos o γ da equação de demanda do estágio superior, equação (2.5), como 1. Enquanto que a estimação do LAIDS produz diversos coeficientes de preço, o PCAIDS produz apenas um. No PCAIDS, todos os outros parâmetros se tornam funções que dependem desse coeficiente único e dos market shares (Coloma 2006). Utilizou-se então os coeficientes das tabelas 2.10 e 2.11 e as equações (2.12) e (2.13) para chegar nas elasticidades apresentadas nas tabelas 5.11 e 5.12. Ao analisarmos as elasticidades provenientes de coeficientes obtidos por SUR, verificamos que as diferenças, em valores absolutos, foram bastante amplas. Temos a elasticidade própria do produto 1 com valor de -2,491 para modelo log-linear e -0,034 para o modelo logit: uma diferença absoluta de 2,457. Segundo a teoria econômica para produtos substitutos e normais, os consumidores demandariam menos de um produto quando

seu preço aumentasse e passariam a demandar mais do seu substituto, gerando assim elasticidades com sinais própria negativas e cruzadas positivas. As elasticidades encontradas com nossos dados apresentaram, portanto, os sinais esperados quando provenientes de coeficientes significativos.

Quando verificamos as elasticidades próprias e cruzadas calculadas a partir dos coeficientes dos modelos estimados por MQ3E, as diferenças de valores, de uma mesma elasticidades mas variando o modelo, seguem discrepantes. Tomando como exemplo novamente a elasticidade própria do produto 1, temos o valor de -1,524 para o modelo log-linear e -0,618 para o modelo logit: uma diferença em valores absolutos de 0,906. As elasticidades apresentaram os sinais de acordo com a teoria econômica em todas aquelas que foram calculadas a partir de coeficientes significativos, com a exceção da elasticidade entre o produto 2 e 3.

Tabela 5.11: Elasticidades calculadas a partir dos coeficientes obtidos por SUR.

	1	2	3
Log-linear			
1	-2.491*	1.339*	0.151*
2	1.088*	-2.649*	-0.018
3	0.736*	1.032*	-1.373*
LAIDS			
1	-1.190*	1.955*	0.302*
2	2.169*	-1.419*	0.244
3	0.302*	0.244	-14.317
PCAIDS			
1	-0.034*	0.839*	0.197*
2	0.945*	-0.140*	0.197*
3	0.945*	0.839*	-0.782*
Logit			
1	-2.440*	0.077*	0.009*
2	0.086*	-2.413*	0.009*
3	0.086*	0.077*	-1.647*

Fonte: Autoria própria.

* indica as elasticidades calculadas a partir de coeficientes significativos.

Tabela 5.12: Elasticidades calculadas a partir dos coeficientes obtidos por MQ3E.

	1	2	3
Log-linear			
1	-1.514*	0.617*	-0.009
2	-0.024	-1.796*	-0.250
3	0.660	1.252	0.566*
LAIDS			
1	-0.674*	1.590*	0.173
2	1.770*	-0.676*	-0.051*
3	0.173	-0.051*	-8.471
PCAIDS			
1	-0.257	1.020	0.240
2	1.150	-0.387	0.240
3	1.150	1.020	-1.168
Logit			
1	-0.618*	0.019*	0.002*
2	0.022*	-0.611*	0.002*
3	0.022*	0.019*	-0.417*

Fonte: Autoria própria.

* indica as elasticidades calculadas a partir de coeficientes significativos.

5.4 Resultados dos testes

5.4.1 Modelos Aninhados

As restrições do PCAIDS foram testadas conjuntamente utilizando um teste do tipo de Wald, já que as restrições implicadas pelo modelo PCAIDS são equações produzidas a partir do modelo LAIDS que podem ser linearizadas da seguinte forma: $\gamma_{jj}(1 - s_k)s_k - s_j(1 - s_j)\gamma_{kk} = 0$ e $\gamma_{jk}(1 - s_k) + s_j\gamma_{kk} = 0$. O teste foi aplicado conjuntamente com as três restrições - restrição 1 (equações (3.1)), 2 ((3.2)) e 3 ((3.3)) - para cada um dos três produtos levando em consideração o número de observações, o número de equações do sistema, o número de parâmetros a serem estimados. O resultado dos testes das restrições do modelo PCAIDS a partir dos coeficientes encontrados pela estimação por SUR, tabela 5.13, foi um estatística qui-quadrada com 4 graus de liberdade a 417,880 e valor da probabilidade de 0,000. Podemos rejeitar, então, o modelo PCAIDS a um nível de significância de 1%.

Tabela 5.13: Testes das restrições do modelo PCAIDS a partir dos coeficientes encontrados pela estimação por SUR.

G.L.	Estat.	Valor-p
5	417.880	0.000

Foi realizado um teste tipo de Wald conjunto para as restrições (respectivamente as equações (3.1), (3.2) e (3.3))

O mesmo resultado é encontrado para os testes das restrições do modelo PCAIDS a partir dos coeficientes encontrados pela estimação por MQ3E, tabela 5.14. Com uma estatística qui-quadrada com 5 graus de liberdade, estatística 66,410 e p-valor de 0,000, podemos rejeitar o modelo PCAIDS a qualquer nível de significância.

Como nos dois casos os testes das restrições dizem que o LAIDS seria melhor para explicar a curva de demanda em comparação ao PCAIDS, neste trabalho, não

Tabela 5.14: Testes das restrições do modelo PCAIDS a partir dos coeficientes encontrados pela estimação por MQ3E.

G.L	Estat.	Valor-p
5	66.410	0.000

Foi realizado um teste tipo de Wald conjunto para as restrições (respectivamente as equações (3.1), (3.2) e (3.3))

iremos testar o modelo PCAIDS com os demais modelos.

5.4.2 Modelos Não Aninhados

Na tabela 5.15, apresentamos os resultados dos testes PE e BEM para os modelos não aninhados, com variáveis dependentes iguais ou diferentes, e estimados sem o auxílio de uma variável instrumental. Considerando um nível de significância de 5%, tanto o teste PE quanto o BEM aceitaram o modelo LAIDS com p-valor de 0,000 e rejeitaram o modelo log-linear aos p-valores de 0,550 e 0,211, respectivamente. Nas demais comparações, nenhum teste foi capaz de indicar um modelo como superior, sendo todas as hipóteses rejeitadas a um p-valor de 0,00.

Tabela 5.15: Testes PE e BEM sem auxílio de variável instrumental.

	TestePE:Estat	TestePE:Valor-p	TesteBEM:Estat	TesteBEM:Valor-p
Log-linear/Logit	17.148	0.001	17.148	0.001
Logit/Log-linear	12360.636	0.000	606.192	0.000
Log-linear/LAIDS	167.323	0.000	176.338	0.000
LAIDS/Log-linear	2.111	0.550	4.514	0.211
Logit/LAIDS	642.720	0.000	393.991	0.000
LAIDS/Logit	26.510	0.000	70.618	0.000

Fonte: Autoria própria.

Os resultados dos testes PE e BEM realizados para os modelos não aninhados, com variáveis dependentes iguais ou diferentes, e estimados com o auxílio de variável instrumental, tabela 5.16, foram similares ao dos testes sem o uso de variável instrumental, conseguindo discriminar apenas entre os modelo LAIDS e log-linear.

Nesse caso, o teste PE, ao considerar o modelo LAIDS como hipótese nula e alternativa o log-linear, resultou em um p-valor muito acima do normal (0.999), indicando que o teste apresentou algum tipo de erro. O teste BEM, contudo, rejeitou o modelo log-linear, se considerarmos um nível de significância de 5%, com um p-valor de 0,075 para essa mesma situação e aceitou o modelo LAIDS, ao invertermos as hipóteses, com um p-valor de 0,00. Novamente, temos resultados inconclusivos para as demais comparações, cujas hipóteses apresentaram p-valor igual a 0,000, podendo ser rejeitadas, então, a qualquer nível de significância.

Tabela 5.16: Testes PE e BEM com auxílio de variável instrumental

	TestePE:Estat	TestePE:Valor-p	TesteBEM:Estat	TesteBEM:Valor-p
Log-linear/Logit	66.205	0.000	66.205	0.000
Logit/Log-linear	12886.001	0.000	500.208	0.000
Log-linear/LAIDS	52.650	0.000	57.749	0.000
LAIDS/Log-linear	0.021	0.999	6.900	0.075
Logit/LAIDS	430.804	0.000	235.157	0.000
LAIDS/Logit	30.893	0.000	20.448	0.000

Fonte: Autoria própria.

Portanto, a conclusão que temos a partir dos testes é de que o modelo LAIDS seria superior ao modelo log-linear, mas que não seria possível fazer qualquer afirmação relacionada ao modelo Logit. A indicação de que o modelo LAIDS seria mais adequado para explicar a curva de demanda é feita tanto pelo teste PE quanto BEM quando não utilizamos variáveis instrumentais para realizá-los. Contudo, apenas o teste BEM consegue afirmar que o modelo LAIDS seria superior ao modelo log-linear quando utilizamos variáveis instrumentais - o que poderia confirmar o maior poder do teste BEM quando comparado ao PE como mencionado pelo Manera (1991).

CONCLUSÃO

Nesse trabalho aplicamos a abordagem de testes de hipóteses estatísticas para comparar modelos de demanda, sendo eles aninhados, ou não, e com as variáveis dependentes de cada modelo sendo diferentes ou iguais. Essa metodologia é então aplicada sobre os principais modelos de demanda utilizados pelos órgãos de defesa da concorrência no Brasil.

Para ilustrar os métodos apresentados, uma aplicação empírica é realizada a partir dos dados semanais de uma rede de supermercado americana colhidos durante aproximadamente 390 semanas. O produto escolhido para representar o mercado relevante foi refrigerante de cola.

Os resultados das estimações dos modelos foram elasticidades com sinais condizentes com a teoria econômica quando significativos e, portanto, de acordo com o esperado. As elasticidades apresentaram valores bem discrepantes entre os modelos; reforçando que a escolha arbitrária para análise de um ato concentração não seja apropriada, sendo necessário mais estudos para tentar encontrar um modelo que melhor descreve as decisões dos consumidores.

Dentre os modelos aninhados PCAIDS e LAIDS, as restrições do modelo PCAIDS foram rejeitadas. Portanto, segundo os testes, o modelo LAIDS descreveria melhor as curvas de demanda do que o modelo PCAIDS. Em decorrência a essa conclusão, apenas o modelo LAIDS foi comparado aos demais modelos. Em relação às demais comparações, tanto o teste PE quanto o BEM indicaram que modelo LAIDS seria superior ao modelo log-linear, mas que não seria possível fazer qualquer afirmação relacionada ao modelo Logit.

O trabalho procurou também corrigir um possível problema de endogeneidade através da utilização dos custos como variável instrumental. A presença da endogeneidade das variáveis relacionadas a preços foi testada através dos testes de Hausman e Durbin e confirmada. Os instrumentos também foram analisados e rejeitamos a hipótese de que fossem fracos. O resultado das comparações é o mesmo quando não utilizamos variáveis instrumentais para realizá-los: o modelo LAIDS seria superior ao PCAIDS e ao log-linear. Contudo, apenas o teste BEM conseguiu discriminar entre o modelo LAIDS e o log-linear, confirmando as conclusões de trabalhos dessa literatura que consideraram o teste BEM um teste com maior poder do que o teste PE.

Por fim, precisamos chamar atenção para algumas limitações do trabalho. Não houve correção para possíveis fontes de má especificação como autocorrelação e heterocedasticidade nem em uma investigação se elas ocorrem, ou não, mas seria interessante um estudo nesse sentido.

REFERÊNCIAS

- [1] ATKINSON, A. C. A method for discriminating between models. **Journal of the Royal Statistical Society. Series B (Methodological)**, v.32, n.3, p.323–353, 1970.
- [2] BERA, A. K.; MCALEER, M. Nested and non-nested procedures for testing linear and log-linear regression models. **Sankhyā: The Indian Journal of Statistics, Series B**, v.51, n.2, p.212–224, 1989.
- [3] BOMFIM, R. d. C. Avaliação de atos de concentração no segmento de motos scooters: uma simulação com modelo pcaids. **Dissertação (mestrado) - Universidade Federal do Paraná, Setor de Ciências Sociais Aplicadas, Programa de Pós-Graduação em Desenvolvimento Econômico.**, , f.63, 2016.
- [4] COLOMA, G. Econometric estimation of PCAIDS models. **Empirical Economics**, v.31, n.3, p.587–599, 2006.
- [5] COX, D. R. Tests of separate families of hypotheses. In: BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY. **Proceedings. . .**, 1961. v.1, p.105–123.
- [6] COX, D. R. Further results on tests of separate families of hypotheses. **Journal of the Royal Statistical Society. Series B (Methodological)**, v.24, n.2, p.406–424, 1962.
- [7] CROOKE, P. et al. Effects of assumed demand form on simulated postmerger equilibria. **Review of Industrial Organization**, v.15, n.3, p.205–217, 1999.
- [8] DAMEUS, A. et al. AIDS versus the Rotterdam demand system: a cox test with parametric bootstrap. **Journal of Agricultural and Resource Economics**, v.27, n.2, p.335–347, 2002.

- [9] DASTOOR, N. K. Some aspects of testing non-nested hypotheses. **Journal of Econometrics**, v.21, n.2, p.213–228, 1983.
- [10] DAVIDSON, R.; MACKINNON, J. G. Several tests for model specification in the presence of alternative hypotheses. **Econometrica: Journal of the Econometric Society**, v.49, n.3, p.781–793, 1981.
- [11] DAVIDSON, R.; MACKINNON, J. G. Some non-nested hypothesis tests and the relations among them. **The Review of Economic Studies**, v.49, n.4, p.551–565, 1982.
- [12] DAVIDSON, R.; MACKINNON, J. G. **Econometric theory and methods**. : Oxford University Press New York, 2004. v.5th.
- [13] DAVIDSON, R.; MACKINNON, J. G. et al. Estimation and inference in econometrics. , 1993.
- [14] DEATON, A.; MUELLBAUER, J. An almost ideal demand system. **The American economic review**, v.70, n.3, p.312–326, 1980.
- [15] DEATON, A. S. Model selection procedures, or does the consumption function exist? **Evaluating the reliability of macroeconomic models**, v.5, p.43–65, 1982.
- [16] GRUPO DE TRABALHO DE MÉTODOS EM ECONOMIA (GTME) / DEPARTAMENTO DE ESTUDOS ECONÔMICOS (DEE). **Análise de Demanda**. Documento de Trabalho No. 001/09. Brasília, Dezembro de 2009.
- [17] DESOUSA, S. A. Antitrust mixed logit model. In: MEETING OF THE BRAZILIAN ECONOMETRIC SOCIETY, 33. **Anais. . .**, 2011.
- [18] EPSTEIN, R. J.; RUBINFELD, D. L. Merger simulation: a simplified approach with new applications. **Antitrust LJ**, v.69, p.883, 2001.

- [19] GODFREY, L. G. Testing non-nested models after estimation by instrumental variables or least squares. **Econometrica: Journal of the Econometric Society**, v.51, n.2, p.355–365, 1983.
- [20] GREENE, W. H. **Econometric analysis.**: Pearson Education India, 2003.
- [21] GUJARATI, D. N. **Basic econometrics.**: Tata McGraw-Hill Education, 2009.
- [22] HUSE, C.; SALVO, A. , Estimação e identificação de demanda e de oferta. In Fiuza, E. P. S. Motta, R. S., editors, **Métodos Quantitativos em Defesa da Concorrência e Regulação Econômica**. IPEA, Rio de Janeiro, 2006.
- [23] KENNEDY, P. **A guide to econometrics.**: MIT press, 2003.
- [24] LAFRANCE, J. T. The Silence bleating! of the lambdas: comment. **American Journal of Agricultural Economics**, v.80, n.1, p.221–230, 1998.
- [25] LANCASTER, K. J. A new approach to consumer theory. **Journal of political economy**, v.74, n.2, p.132–157, 1966.
- [26] MACKINNON, J. G.; MAGEE, L. Transforming the dependent variable in regression models. **International Economic Review**, v.31, n.2, p.315–339, 1990.
- [27] MACKINNON, J. G.; WHITE, H.; DAVIDSON, R. Tests for model specification in the presence of alternative hypotheses: some further results. **Journal of Econometrics**, v.21, n.1, p.53–70, 1983.
- [28] MADDALA, G. S.; LAHIRI, K. **Introduction to econometrics.**: Macmillan New York, 1992.
- [29] MANERA, M. Multivariate non-nested tests for comparing the empirical performance of alternative factor demand systems. **Giornale degli economisti e annali di economia**, v.50, n.9/12, p.441–474, 1991.

- [30] MANERA, M.; MCALEER, M. Testing Multiple Non-Nested Factor Demand Systems. **Bulletin of Economic Research**, v.57, n.1, p.37–66, 2005.
- [31] MCALEER, M. The significance of testing empirical non-nested models. **Journal of Econometrics**, v.67, n.1, p.149–171, 1995.
- [32] MIZON, G. E.; RICHARD, J.-F. The encompassing principle and its application to testing non-nested hypotheses. **Econometrica: Journal of the Econometric Society**, v.54, n.3, p.657–678, 1986.
- [33] PESARAN, M. H. On the general problem of model selection. **The Review of Economic Studies**, , v.41, n.2, p.153–171, 1974.
- [34] PESARAN, M. H. Non-nested hypotheses. In: **Econometrics.**: Palgrave Macmillan UK, 1990. p.167–173.
- [35] PESARAN, M. H.; DEATON, A. S. Testing non-nested nonlinear regression models. **Econometrica: Journal of the Econometric Society**, v.46, n.3, p.677–694, 1978.
- [36] PESARAN, M. H.; WEEKS, M. Non-nested hypothesis testing: an overview. **A companion to theoretical econometrics**, p.279–309, 2001.
- [37] STONE, R. **The measurement of consumers' expenditure and behaviour in the United Kingdom, 1920-1938**. v.1.
- [38] TRAIN, K. E. **Discrete choice methods with simulation.**: Cambridge university press, 2009.
- [39] ULLOA, M. R. D.; PESARAN, M. H. Nonnested Hypotheses. , 2006.
- [40] VUONG, Q. H. Likelihood ratio tests for model selection and non-nested hypotheses. **Econometrica: Journal of the Econometric Society**, v.41, n.2, p.307–333, 1989.

- [41] ZELLNER, A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. **Journal of the American statistical Association**, , v.57, n.298, p.348–368, 1962.

APÊNDICE A TESTE DE ENDOGENEIDADE

Tabela A.1: Teste de Endogeneidade

	GL	Estat	p-valor
Log-linear			
Durbin_(score)_chi2()	3	9.608	0.022
Wu-Hausman_F()	381	3.216	0.023
LAIDS			
Durbin_(score)_chi2()	3	25.728	0.000
Wu-Hausman_F()	381	8.995	0.000
Logit			
Durbin_(score)_chi2()	1	5.412	0.020
Wu-Hausman_F()	386	5.432	0.020

Fonte: Autoria própria.

Para verificar a existência de endogeneidade, foram realizados os testes de Durbin e Wu-Hausman. Considerando como hipótese nula as variáveis sendo exógenas, conforme os valores da tabela A.1, podemos rejeitar a hipótese nula a um nível de significância de 5% para os três modelos.

APÊNDICE B TESTE DOS INSTRUMENTOS

Tabela B.1: Teste dos Instrumentos

	GL	Minimum_eigenvalue_statistic
Log-linear	3	17.195
LAIDS	3	17.467
Logit	1	395.688

Fonte: Autoria própria.

Para verificar se um instrumento é fraco em um modelo com múltiplas variáveis endógenas, como é o nosso caso, utilizamos a estatística de Cragg-Donald, com distribuição F. Considerando como hipótese nula que os instrumentos sejam fracos, segundo os valores das estatística apresentados na Tabela B.1, podemos rejeitar esta hipótese para os três modelos.