UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DE RIBEIRÃO PRETO
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA – ÁREA: ECONOMIA APLICADA

CAMILA DE MOURA VOGT

Competição e dispersão de preços de gasolina e etanol no mercado brasileiro de combustíveis: evidências do município de São Paulo

ORIENTADOR: PROF. DR. CLÁUDIO RIBEIRO DE LUCINDA

RIBEIRÃO PRETO
2014

Prof. Dr. Marco Antonio Zago
Reitor da Universidade de São Paulo

Prof. Dr. Sigismundo Bialoskoski Neto
Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Prof. Dr. Sérgio Kannebley Junior
Chefe do Departamento de Economia

Profa. Dra. Elaine Toldo Pazello
Coordenador do Programa de Pós-Graduação em Economia – Área: Economia Aplicada

CAMILA DE MOURA VOGT

Competição e dispersão de preços de gasolina e etanol no mercado brasileiro de combustíveis: evidências do município de São Paulo

Dissertação apresentada ao Programa de Pós-Graduação em Economia – Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo para obtenção do título de Mestre em Ciências. Versão Corrigida. A original encontra-se disponível na FEA-RP/USP

ORIENTADOR: PROF. DR. CLÁUDIO RIBEIRO DE LUCINDA

RIBEIRÃO PRETO
2014

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Vogt, Camila de Moura
Competição e dispersão de preços de gasolina e etanol no mercado brasileiro de combustíveis: Evidências do Município de São Paulo, 2014.
59 p.: il.; 30 cm

Dissertação de Mestrado, apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto/USP.
Orientador: Ribeiro de Lucinda, Cláudio.

1. Etanol. 2. Gasolina. 3. Concorrência. 4. Dispersão de Preços. 5. Custo de procura.

FOLHA DE APROVAÇÃO

Nome: Camila de Moura Vogt
Título: Competição e dispersão de preços de gasolina e etanol no mercado brasileiro de combustíveis: evidências do município de São Paulo.

Dissertação apresentada ao Programa de Pós-Graduação em Economia – Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Aprovado em:

BANCA EXAMINADORA

Prof° Dr. _____

Instituição:_____ Assinatura:_____

Prof° Dr. _____

Instituição:_____ Assinatura:_____

Prof° Dr. _____

Instituição:_____ Assinatura:_____

RIBEIRÃO PRETO

2014

AGRADECIMENTOS

Agradeço ao Programa de Pós-Graduação em Economia e à Capes pela bolsa de mestrado de 12 meses.

Ao meu orientador Cláudio Lucinda pela atenção, paciência e todas as lições nesse último ano.

À Agência Nacional do Petróleo, Gás Natural e Biocombustíveis -ANP pelos dados fornecidos, essenciais para a execução desse trabalho.

Aos professores Bruno Ledo e Rodrigo Moita, pelas contribuições à esse trabalho durante a banca de qualificação.

À Adam Ozimek e Stefan Bernhard pelo auxílio no uso do Google Maps API.

Às colegas de mestrado Graziella Magalhães e Thaís Barcellos e à todos os amigos que apoiaram de alguma forma para concretização desse trabalho.

À minha família, meus pais João e Ossênia, e meus amados irmãos Helena, Maurício, Otávio, Alice e o pequeno João. Ao querido Leonardo Kerber, por todo apoio e carinho.

"[...] *knowledge is power. And yet it occupies a slum dwelling in the town of economics*"

George J. Stigler (1961)

RESUMO

VOGT, C. M. **Competição e dispersão de preços de gasolina e etanol no mercado brasileiro de combustíveis: evidências do município de São Paulo. 2014. 59 f.** Dissertação de Mestrado. Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo.

O fenômeno da dispersão de preços é mais comum do que se supõe a partir da microeconomia básica. O presente estudo busca avaliar a importância dos custos de procura na dispersão de preços de gasolina e etanol no município de São Paulo. Para tanto, são utilizados como referenciais teóricos modelos de dispersão e estudos de estruturas de mercado. A parte empírica contempla duas etapas, na primeira é estimado qual o mercado relevante para o varejo de combustíveis e, posteriormente, a descrição da dispersão na área. Os dados dos preços de varejo são coletados semanalmente desde o ano de 2001 e a série histórica foi cedida pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Os resultados das estimações nos mostram que, apesar da dispersão, possivelmente existe maior relação com a heterogeneidade dos estabelecimentos do que com a busca do consumidor.

Palavras Chave: Etanol, Gasolina, Concorrência, Dispersão de Preços, Custo de Procura.

ABSTRACT

VOGT, C. M. **Competition and price dispersion of gasoline and ethanol in the Brazilian fuel market: evidence from São Paulo city. 2014. 59 p.** Master Dissertation. Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo.

The price dispersion phenomenon is more common than might be supposed from basic microeconomics. This study aim to evaluate the importance of search costs in the dispersion of fuel prices in São Paulo city. As theoretical reference, we use models of price dispersion and market structures studies. The empirical part has two stages, first is estimated the relevant retail fuel market, and then we describe the dispersion in the area. The retail prices data of gasoline and ethanol have been collected since 2001 and the time series was provided by the Agência Nacional do Petróleo, Gás Natural e Biocombustíveis - ANP. The results shown that although there is dispersion it is more closely related to the establishments heterogeneity, than with consumer search.

Key words : Ethanol Gasoline Competition, Price Dispersion, Search Cost.

LISTA DE GRÁFICOS

Gráfico 01 - Dispersão de preços gasolina e etanol - Janeiro a julho 2012. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor 15

Gráfico 02 - Proporção de consumidores informados μ. Chandra e Tappata (2011) 22

Gráfico 03 - Proporção de consumidores informados $\boldsymbol{\mu}$ e número de firmas. Chandra e Tappata (2011) 23

Gráfico 04 - Número de Observações Analisadas por Período. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor.....32

Gráfico 05 - Dispersão de Preços de Venda de Gasolina e Etanol 01/2004. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor 33

Gráfico 06 - Histograma Dispersão de Preços de Venda de Gasolina e Etanol 01/2004. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor ..33

Gráfico 07 - Histograma Dispersão de Preços de Venda de Gasolina e Etanol 01/2004. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor ..33

Gráfico 08 - Dispersão de Preços de Venda de Gasolina e Etanol 01/2009. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor 34

Gráfico 09 - Histograma Dispersão de Preços de Venda de Gasolina e Etanol 01/2009. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor ..34

Gráfico 10 - Histograma Dispersão de Preços de Venda de Gasolina e Etanol 01/2009. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor ..34

Gráfico 11 - Dispersão de Preços de Venda de Gasolina e Etanol 01/2011. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor 35

Gráfico 12 - Histograma de Preços de Venda de Gasolina e Etanol 01/2011. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor 35

Gráfico 13 - Histograma de Preços de Venda de Gasolina e Etanol 01/2011. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor 35

Gráfico 14 – *Ranking reverso* gasolina (5,0 a 6,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor.....44

Gráfico 15 – *Ranking reverso* etanol (5,0 a 6,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor.....45

LISTA DE TABELAS E QUADROS

Tabela 01 – Raios de competição. Elaboração: Autor 37

Tabela 02 – Estimações Raios de Proximidade – Mínimos Quadrados Ordinários e Efeitos Fixos 41

Tabela 03 – Estimações Raios de Proximidade com Variáveis Instrumentais 42

Tabela 04 – Resultados *rank* reverso. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor 46

Quadro 01 – Resultados estimações *rank* reverso. Vizinhos *versus* "Esquina". Elaboração: Autor 47

SUMÁRIO

1. INTRODUÇÃO 13

2. REVISÃO DE LITERATURA 16

2.1 Estudos sobre dispersão de preços 16

2.2 Modelos teóricos de dispersão de preços 17

2.3 Estudos empíricos de dispersão de preços 24

3. METODOLOGIA 30

3.1 Base de dados 30

3.1.1 Levantamento de preços ANP 30

3.1.2 Georreferenciamento 35

3.2 Modelo empírico 36

4. RESULTADOS E DISCUSSÃO 40

4.1 Raios de competição 40

4.2 Dispersão temporal de preços 43

5. CONCLUSÃO 48

6. BIBLIOGRAFIA 51

APÊNDICE 53

ANEXOS 59

1. INTRODUÇÃO

Quando pensamos em preços de bens idênticos, a intuição nos leva a concluir que mercadorias com as mesmas características deveriam ser comercializadas pelo mesmo valor, salvo os custos de transporte, tarifas e outras singularidades do mercado. Ao observarmos a realidade, entretanto, nos deparamos com situações opostas à intuição natural. É comum, principalmente nos mercados de varejo, encontrar mercadorias exatamente iguais vendidas por valores diferentes, mesmo em estabelecimentos vizinhos.

Stigler (1961), acrescentou uma variável fundamental para a explicação do porquê de bens considerados idênticos serem comercializados por valores diferentes: a informação. Para o autor, o custo de procura pela melhor oferta afeta diretamente a estratégia de venda das firmas, e contribui para a dispersão de preços no mercado.

Mensurar como isso irá afetar as estratégias das firmas e consumidores é, no entanto, um campo pouco explorado. Por se tratar de uma característica não observável, o custo de procura de cada consumidor irá depender de quais são as fontes de informação de preços, e como e com qual intensidade o acesso a elas irá influenciar na dispersão. Além do mais, é necessário limitar qual o espaço físico do mercado que é suscetível à procura dos consumidores.

Esse trabalho entende a importância do estudo do custo de procura e da informação no mercado brasileiro, e visa estudar como essas variáveis irão contribuir para o ambiente competitivo através da dispersão de preços. Para isso, foram escolhidos como objetos de análise os postos de combustíveis no município de São Paulo. Essa opção foi feita principalmente por justificativas que se aproximam das utilizadas no estudo de Chandra e Tappata (2011)[1]: os custos marginais dos estabelecimentos serem muito próximos; os preços dos produtos estarem acessíveis facilmente a todos os consumidores; o munícipio de São Paulo apresentar características de mercado que facilitam o entendimento da dispersão; e a disponibilidade de dados semanais dos preços e localização geográfica do varejo de combustíveis brasileiro.

Acrescenta-se ainda à escolha o fato desses estabelecimentos no Brasil responderem pelo abastecimento de uma frota de certa de 34,8 milhões de veículos[2] que garantem o transporte e logística do país. Portanto, um mercado de extrema importância que motiva o entendimento aprofundado sobre sua estrutura de preços.

---

[1]Chandra e Tappata (2011) ao analisar a dispersão de preços nos postos de combustíveis dos EUA cita as seguintes motivações para a análise desse mercado: i) demanda inelástica; ii) não há a possibilidade de armazenamento; iii) custos marginais parecidos, iv) não há restrição de capacidade do produto, v) preços facilmente disponíveis aos consumidores.

[2]Levantamento da frota circulante brasileira em 2011.

No estudo são analisados os preços dos combustíveis gasolina comum e etanol. Essa opção leva em consideração um fator específico ao caso do mercado brasileiro, a questão da ampla disponibilidade dos dois tipos de produtos. A grande importância do carro *flex*[3] no mercado também levanta o questionamento sobre a relação de consumo entre gasolina e etanol, e de que forma e com qual intensidade ela irá afetar os preços e comportamento dos consumidores.

Toda a base de dados do estudo foi fornecida pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP, que possui um sistema de coleta semanal dos preços em todo o Brasil desde 2001. Os endereços dos postos coletados no município de São Paulo foram georreferenciados, e foi elaborado um banco de dados com a distância de cada um dos estabelecimentos em relação a seus vizinhos.

Utilizando os dados dos preços, e a localização dos estabelecimentos, é feita uma estimativa a fim de mensurar qual a área de mercado relevante dos postos de combustíveis para a análise da dispersão. Utilizando como base o trabalho de Davis (2005), que analisa impacto da concorrência a partir das informações correspondentes a estruturas de mercado, verificamos uma área mínima para a análise do grau de dispersão.

Como referencial teórico e empírico sobre dispersão foi utilizado o artigo de Chandra e Tappata (2011), que constrói um modelo que permite relações entre o número de consumidores informados e dispersão de preços. Esse modelo é a base para o estudo empírico do comportamento dos preços de venda de varejo de combustíveis nos Estados Unidos. Também é construído um índice que considera as alterações de preços ao longo do tempo a fim de medir a dispersão gerada pelo custo de procura. O fato dos preços estarem expostos em frente aos estabelecimentos, permite ainda criar uma amostra de controle, onde o custo de informação é zero, permitindo a comparação com vizinhanças distantes onde há custos de informação.

Assim, após verificar qual o mercado relevante para o estudo de dispersão, e traçadas as vizinhanças dos postos de combustível, é calculado o índice de dispersão por proximidades. Empregando o modelo teórico e os resultados do índice para a amostra de controle, é possível verificar qual a área ótima de procura considerando a dispersão de preços e qual a influência da informação no mercado.

Com relação a base de dados, uma evidência preliminar da dispersão de preços pode ser encontrada no gráfico 01 onde é possível ver os preços de etanol e gasolina em um mesmo

[3]Veículos movidos a combustíveis duplos, gasolina ou etanol.

período de 2012. Temos no eixo horizontal os preços de gasolina, e no eixo vertical os de etanol, e cada ponto corresponde a uma combinação de preços praticada por um posto de gasolina no período de coleta referente aos primeiros sete meses do ano de 2012. A reta de valores "relação técnica" e a linha de tendência com intervalo de confiança 95% permite ter noção sobre a dispersão. Em muitos casos a relação entre estes preços está bastante fora do que se esperaria, dado o conteúdo energético dos dois combustíveis: o etanol tem cerca de 30% menos eficiência energética que a gasolina[4].

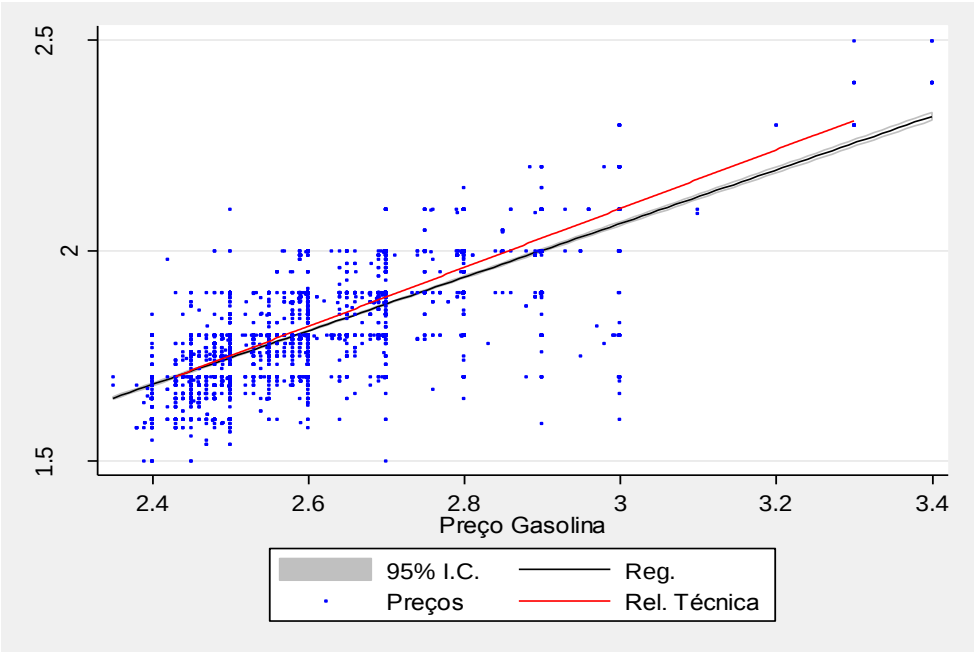


Gráfico 01 - Dispersão de preços gasolina e etanol - Janeiro a julho 2012. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor

Dessa forma, considerando as justificativas apresentadas, esse trabalho está dividido em quatro capítulos. No primeiro, é feita a revisão da literatura de dispersão de preços e discutidos modelos teóricos e empíricos sobre custo de procura e estruturas de mercado. No segundo é apresentada a metodologia escolhida para o estudo de dispersão, baseada nos modelos de Davis (2006) e Chandra e Tappata (2011), além de apresentados os dados e ferramentas de estimação. Já no terceiro capítulo são apresentados os resultados e discutido quais são as dimensões geográficas da competição no mercado, e como o custo de procura impactou nos preços dos postos de combustíveis. Finalmente, na última parte do trabalho será feita a conclusão com base nos resultados e modelos teóricos.

[4]Disponível em: http://www.inee.org.br/etanol_hidratado.asp?Cat=etanol. Acesso: 28/03/2014.

2. REVISÃO DE LITERATURA

2.1 Estudos sobre dispersão de preços

Estabelecer critérios sobre o porquê de algumas empresas em equilíbrio de mercado cobrarem preços diferentes é uma questão importante na microeconomia. Varian (1980), em seu estudo sobre equilíbrio focando no papel das estratégias mistas de preços, afirma que é um desafio para a teoria econômica explicar como a dispersão pode persistir em mercados com agentes racionais.

Habitualmente a dispersão é atribuída a quatro fenômenos na literatura de microeconomia: a diferenciação do produto, a discriminação de preços, choques de oferta e a busca do consumidor.

Para o mercado de postos de combustíveis em que o produto é homogêneo e os custos marginais de revenda e armazenamento são similares, um ponto de partida para o estudo da dispersão é o artigo de Stigler (1961), sobre o custo de informação e busca do consumidor. Nele o autor acrescenta o conceito de custo de procura (ou, em inglês, *search cost*), que aponta a dispersão de preços como um resultado natural da existência de consumidores com informação imperfeita. O autor inclui no custo do produto o custo da informação, que é considerado como o valor necessário para que os consumidores sejam informados dos diversos preços praticados. Dessa forma, dependendo de como se comporta esse custo nos mercados, os preços serão diferentes entre as firmas, e isso irá ultrapassar as características de custo e produto.

O papel da informação na dispersão é revisitado no trabalho de Salop e Stigltz (1977), em que temos consumidores denominados “informados” e “desinformados”. Neste caso, os consumidores “informados” irão sempre atrás do menor preço, pois possuem a informação sobre todos os preços no mercado. Os “desinformados”, por outro lado, irão possivelmente aceitar um preço mais alto, já que não possuem conhecimento sobre como eles se distribuem. Já as firmas irão ter curvas de custo idênticas, e se comportar como agentes formadores de preços monopolisticamente competitivos.

Para Salop e Stigltz (1977), em equilíbrio, uma fração das firmas cobra o preço de monopólio e outra fração pratica o preço competitivo. As firmas com maior preço terão como clientela somente os consumidores não informados, desde que seja suficiente para manter os lucros. O equilíbrio é encontrado quando uma parcela adota o preço mínimo e outra parcela o preço máximo, o que caracteriza modelos de dispersão de preço espacial.

Apesar dos estudos iniciais da teoria de preços somente considerarem a dispersão espacial, ou seja, diferentes preços entre diferentes firmas, o fato de que experiências repetidas passam a acrescentar conhecimento sobre o mercado traz a classificação da dispersão temporal de preços e constitui a base teórica dos estudos de dispersão.

Varian (1980) explica que uma vez que os consumidores aprendam com suas experiências, a persistência do preço mais alto não faria sentido, introduzindo então os modelos de dispersão temporal.

A dispersão temporal, assim como a espacial, é frequentemente observada nos mercados, principalmente nas vendas do varejo. São comuns promoções, ou alterações de preços com duração definida. Dessa forma, é necessário considerar além da dispersão entre os estabelecimentos, a dispersão ao longo do tempo.

O desafio dos estudos consiste em explicar qual seria o preço de equilíbrio de acordo com a informação disponível aos consumidores. Os modelos teóricos se tornam a principal ferramenta na tentativa de estabelecer essa variável, e explicar o impacto da informação e do custo de procura no comportamento dos preços.

2.2 Modelos teóricos de dispersão de preços

De acordo com Marvel (1976), um modelo completo de dispersão de preços deve incluir especificações sobre como as firmas se comportam respondendo a mudanças no nível de informações dos consumidores. A teoria precisaria ser capaz de avaliar o papel da informação sobre a dispersão nos preços de diferentes produtos, e a partir disso poder guiar os testes empíricos.

Os modelos de equilíbrio consideram a teoria da informação em dispersão de preços a partir de estratégias puras e estratégias mistas, com livre entrada e saída de empresas. No primeiro caso se considera que exista um equilíbrio de Nash, ou seja, nenhuma firma irá lucrar mais mudando sua estratégia de preços unilateralmente. O que não ocorre no caso de estratégias mistas. Nessas duas linhas destacam-se os modelos de Salop e Stiglitz (1977) e Diamond (1971), respectivamente.

Também se observa na literatura a busca não sequencial e busca sequencial. Para o modelo não sequencial de dispersão (VARIAN, 1980) o consumidor decide entre comprar da loja com o menor preço, incorrendo no custo s de informação, ou comprar a partir de uma busca aleatória sem nenhum custo. O consumidor desinformado irá procurar aleatoriamente, e sendo

o preço do produto em questão menor do que o valor de reserva r, ele irá adquirir o produto. Já o consumidor informado por conhecer a distribuição de preços irá buscar o produto de menor valor. A busca sequencial (STAHL, 1989) implica que o consumidor observa vários preços (sem qualquer custo) e, então, decide entre a busca de um preço mais baixo (a custo s) ou comprar no preço mínimo observado. Na busca sequencial, consideramos que não estamos lidando com um ambiente com um tamanho fixo, e sim uma procura de forma sequencial, onde o número de buscas é uma variável aleatória com distribuição geométrica.

Conforme Baye et al. (2006) a principal vantagem da busca sequencial é a economia de custo de procura. O consumidor pesa a expectativa de benefício e custos de nova procura logo após cada nova informação encontrada. Se um preço razoável é obtido no início, então os ganhos esperados em se continuar a busca são pequenos, e, portanto, não há a necessidade de continuá-la. Já a vantagem em se ter a busca dentro de uma amostra fixa, é a rapidez com que se obtém a informação.

O modelo não-sequencial de Varian (1980), acrescenta a ideia de dispersão temporal, pois considera que os consumidores aprendem com suas buscas o que irá influenciar na distribuição de preços ao longo do tempo. Acrescentado a isso, o modelo usa a possibilidade de aleatoriedade na busca de preços, com base em Shilony (1977).

Varian (1980) descreve como um exemplo de fonte de informação um jornal semanal com a divulgação de preços, quem tiver acesso a ele necessariamente vai saber a distribuição temporal dos preços e buscar valores mais baixos. Dessa forma, pode existir um custo para adquirir essa informação, definido como s. Consideramos ainda que o número de informados I, e desinformados M é maior que zero logo, $I > 0$ e $M > 0$. Sendo o número de lojas igual a n, o número de desinformados por loja passa a ser $U = \frac{M}{n}$. Cada loja vai possuir uma função de densidade que indica a probabilidade que ela tem de cobrar o preço p[5].

A estratégia de se cobrar determinado valor vai ser pensada dadas as estratégias das outras firmas e o comportamento dos consumidores, constituindo um equilíbrio de Nash. Como o modelo considera a dispersão temporal, a cada semana as lojas irão de maneira aleatória mudar o seu preço. Intuitivamente a loja que for mais bem sucedida terá $I + U$ consumidores, e a que tiver o maior preço terá somente a parcela U de consumidores. O modelo também considera que as firmas possuem uma função de custo médio idêntica e decrescente. Espera-se que as entradas de firmas ocorram até os lucros econômicos chegarem a zero. Para o autor, a

[5]Varian (1980) utiliza a função de probabilidade $(1 - F(p))^{n-1}$

randomização dos preços entre as firmas é de interesse para a discriminação de preços entre os consumidores “informados” e os “desinformados”. E apesar de seu modelo encontrar um preço de equilíbrio, ele explicita que raramente se observam grandes varejos utilizando valores intermediários para a venda de seus produtos.

Baye et al. (2006) possui algumas conclusões sobre os modelos teóricos de dispersão de preços. Em primeiro lugar, não existe um modelo que irá responder por todos os casos de dispersão. Isto se intensifica quando a acessibilidade aos preços dos produtos é maximizada pelos ambientes on-line. O fato das lojas disponibilizarem as informações de valores de venda de modo irrestrito para os consumidores amplia o acesso à informação e consequentemente a comparação de preços. Nesses ambientes os custos de procura caem a valores muito baixos, o que não necessariamente acaba com a dispersão.

Nesse aspecto se ressalta também que questões subjetivas relacionadas à fidelização do consumidor dificultam a análise. Um exemplo são as promoções de fidelização de grandes redes de varejo que incentivam o cliente a utilizar determinado fornecedor e acumular vantagens de compras. A venda desse fornecedor não irá ocorrer diretamente em função do preço, da qualidade, ou do custo de procura do produto, mas sim de uma característica de diferenciação subjetiva.

Outro ponto trazido pelos autores é o fato de que a distribuição dos preços também está diretamente ligada à interação entre estruturas de mercado, número de participantes, elasticidade da demanda e quesitos de diferenciação. São inúmeras as variáveis, que a partir de um comportamento conjunto, irão permitir conclusões sobre a dispersão de preços, o que reforça a ideia de Baye et al. (2006) sobre não haver um modelo teórico que responda por todos os casos.

A teoria também não se pronuncia sobre as previsões de um modelo que combine diferenciação do produto e busca do consumidor. O que se torna um problema para a separação desses dois efeitos, conforme apontando por Chandra e Tappata (2011). De acordo com os autores, os modelos estáticos comparativos, ou seja, que não avaliam a dispersão ao longo do tempo, normalmente não identificam todo o efeito do custo de procura na construção dos preços.

Chandra e Tappata (2011) propõe um teste sobre os impactos da informação no mercado de postos de combustíveis americano a partir da comparação, ao longo do tempo, dos preços relativos entre os estabelecimentos, utilizando, para isso, grupos de controle para informação imperfeita. O modelo teórico, no entanto, assume uma série de suposições baseadas em modelos

clássicos como Varian (1980). Os autores consideram um mercado homogêneo com n firmas competindo por preços e com o mesmo custo de produção marginal, ou seja, para cada unidade a mais de produto um custo igual a c. Sendo o lado da demanda caracterizado por consumidores com demanda inelástica igual a v.

Uma fração λ com distribuição entre (0,1) tem zero custo de procura, e outra fração denominada "compradores" tem custo de procura diferente de zero. A fração sem custo de procura, "não-compradores", decide o local de compra antes de observarem os preços, comprando de maneira aleatória. Conforme os autores a busca não sequencial é escolhida como modelo predileto para o mercado de combustíveis[6]. A proporção de consumidores que opta pela informação irá definir a intensidade μ de busca desse mercado.

As condições de equilíbrio do modelo serão: i) Para a intensidade[7] $\mu \epsilon [\lambda, 1]$; ii) as estratégias das firma precisam seguir um equilíbrio de Nash em estratégias mistas; iii) a intensidade de procura do mercado $\mu *$ precisa estar consistente com a estratégia de preços das firmas, ou seja, os consumidores comparam os custos com os benefícios da procura e assim antecipam corretamente as estratégias de preços das firmas.

Dessa forma, em cada período, as firmas escolhem suas estratégias simultaneamente conforme função de probabilidade (1)

$$F(p, c, v, n) = 1 - [\frac{(1-\mu)}{\mu n}\frac{(v-p)}{(p-c)}]^{1/(n-1)} \quad (1)$$

Considerando que o preço que maximiza a função é tal que: $p \in \left[p^* = \frac{cn\mu+(1-\mu)v}{1+(n-1)\mu}, v\right]$. Para preços abaixo de p^*, a firma prefere sempre cobrar o preço de monopólio e vender para $(\frac{1-\mu}{n})$ consumidores. A procura será sempre relacionada com a intensidade de competição. Se o número de consumidores informados cresce, o limite será que a firma irá cobrar o custo marginal. No outro extremo, se o número de consumidores informados é zero, ou seja, não há busca, $p = v$ e cada firma se torna um monopolista sobre $\frac{1}{n}$ consumidores.

Os autores propõem a seguinte função não-monotônica de intensidade de procura associada aos ganhos de procura. Em outras palavras quanto o consumidor irá ganhar pela

[6]Conforme os autores Chandra e Tappata (2011) as estimações de modelos de busca sequencial, são similares os de busca não-sequencial.
[7]A intensidade é definida pela proporção de consumidores informados.

diferença entre o intervalo do preço máximo e o preço mínimo cobrado pelas firmas em função da quantidade de informação.

A equação é definida pela sigla GS (*gains from search*):

$$GS = E[p - p_{min}|\mu; c, \mathrm{v}, \mathrm{n}] = \int_{\mathrm{p}^*}^{\mathrm{v}} \mathrm{p}[1 - \mathrm{n}[1 - \mathrm{F}(\mathrm{p}; \mathrm{c}, \mathrm{v}, \mathrm{n})]^{\mathrm{n}-1}]\mathrm{dF}(\mathrm{p}; \mathrm{c}, \mathrm{v}, \mathrm{n}) \quad (02)$$

Para a avaliação durante situações de monopólio ou concorrência perfeita, se considera o preço máximo e preço mínimo, $\mu = 0\ e\ \mu = 1$, ou consumidores completamente informados e completamente não informados.

Como o modelo é não-monotônico, os ganhos de procura são baixos quando muitos ou poucos consumidores decidem comprar, mas assume ganhos altos quando a intensidade de procura é feita de maneira intermediária. O gráfico 02 mostra as medidas de dispersão de preços em função da intensidade da pesquisa. Tanto o intervalo de amostragem (SR) dos ganhos de procura (equação 02) e o desvio-padrão (DP) assemelham-se à forma do GS (função de intensidade de procura).

Conforme o modelo, dispersão por si só não pode ser usada para predizer a intensidade de pesquisa ou nível de competição no mercado. Para isso, devemos definir como a intensidade de busca maximiza GS para um determinado custo de produção e número de firmas. A relação entre a intensidade de busca e dispersão de preços é positiva quando $\mu^* < \hat{\mu}$ e negativo quando $\mu^* > \hat{\mu}$. Assim, intuitivamente, o modelo nos diz que mais consumidores informados geram mais dispersão quando o equilíbrio está abaixo da intensidade de procura, ou seja, há mais compradores informados do que o equilíbrio necessita. De forma oposta, mais consumidores informados geram menos dispersão quando a intensidade de procura está abaixo do equilíbrio.

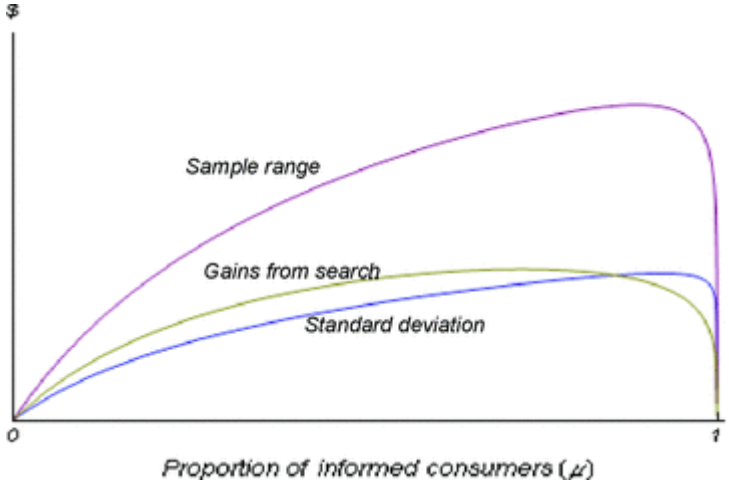


Gráfico 02 - Proporção de consumidores informados $\boldsymbol{\mu}$. Chandra e Tappata (2011)

Para o equilíbrio é necessário ainda que o montante de procura μ^* seja consistente com as estratégias de preços das firmas. Ou seja, cada consumidor antecipa sua estratégia (equação 01), tendo a intensidade de procura como dada, e compara os ganhos de procura (equação 02) com seus gastos.

Os consumidores que buscam aleatoriamente, "não-compradores", constroem seus custos de procura com uma distribuição contínua $G(s_i)$ com $s_i \in [0, \bar{s}]$. No equilíbrio o consumidor indiferente tem um custo de procura $\tilde{s}$ tal que $GS = \tilde{s}$ e a intensidade de procura se torna $\mu * = \lambda + GS(\widetilde{s})(1 - \lambda)$. Os "não-compradores" irão desenhar suas estratégias pesando o ganho de procura.

No gráfico 03 A[8], a partir de uma demonstração dos autores que utilizam uma distribuição Beta para custos, podemos observar o comportamento dos "não-compradores" com relação ao número de firmas e custo marginal de procura. A proporção de consumidores informados é representada pelo eixo horizontal, enquanto os ganhos de procura e custos pelo eixo vertical. As linhas côncavas do gráfico representam os ganhos de procura de consumidores em mercados com 5 e 10 firmas. Cortando as curvas de ganho de procura temos o equilíbrio nesses mercados a partir do custo marginal de consumidor que decide pela busca.

Podemos observar, portanto, que quando o número de firmas aumenta no mercado, o custo de procura de um consumidor indiferente é maior, implicando que quanto mais atomizado o mercado maior a intensidade de procura (Gráfico 03 A e B).

[8]Parâmetros usados pelo autor $\lambda = 0{,}15$, $v = 1$, $c = 0$, $G(s_i) = I[2{,}2]$, sendo I a função Beta.

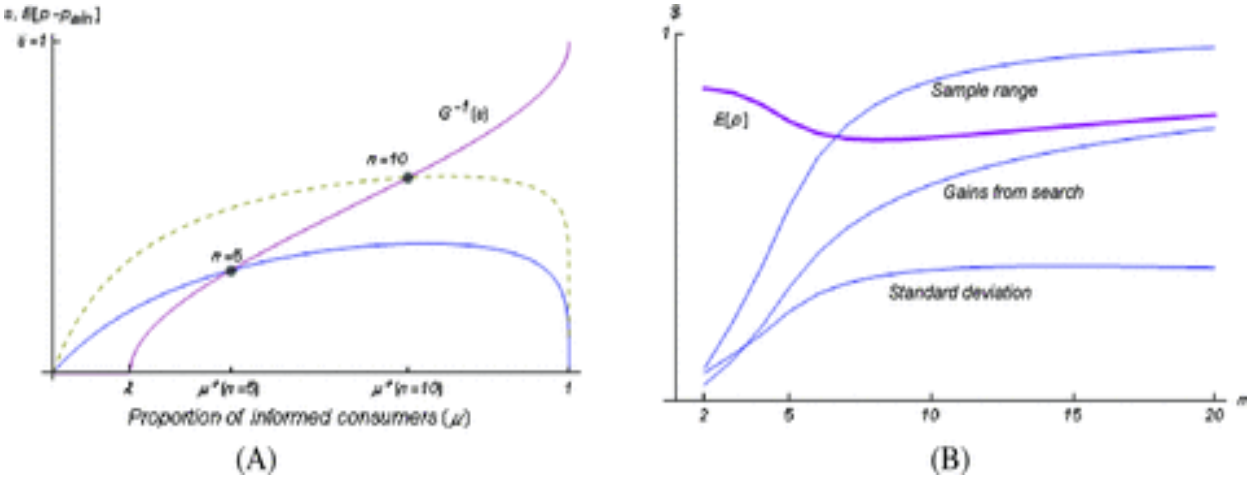


Gráfico 03 - Proporção de consumidores informados $\boldsymbol{\mu}$ e número de firmas. Chandra e Tappata (2011)

Sobre a análise de entrada e saída de firmas, assume-se que a intensidade de procura é fixa ($\mu = \bar{\mu}$). Dessa forma, quando o número de firmas aumenta, os ganhos esperados mudam de duas diferentes formas. A proporção de consumidores não informados captados decresce e a probabilidade de ter o preço mais baixo do mercado decresce também. Considerando que os ganhos de baixar os preços diminuem com uma taxa maior que os ganhos de aumentar os preços, a distribuição de preços irá mudar em direção aos preços mais altos, e a dispersão de preços irá aumentar.

Para ter o efeito completo do número de firmas nos níveis de dispersão é necessário incorporar as reações dos consumidores informados. Considerando que para um dado μ, a dispersão de preços cresce com n, o consumidor marginal terá maiores custos de procura em um mercado com mais firmas. Quando $\mu < \hat{\mu}$, maiores intensidades de busca reforçam o efeito parcial e positivo do número de firmas na dispersão de preços. Quando $\mu > \hat{\mu}$ uma maior intensidade de buscas reduz esse efeito, apesar de nunca o anular.

Em relação ao aumento de consumidores, entretanto, não podemos considerar o efeito total de mais consumidores informados e desinformados no preço médio, já que mais consumidores procurando diminuirão os preços de equilíbrio.

É também importante ressaltar a questão da busca não-sequencial no modelo. Em nenhum momento é considerado que um maior número de firmas aumente os custos de procura dos consumidores. O que ocorre devido a hipótese do consumidor decidir antecipadamente entre comprar da loja com o menor preço com o custo s de procura, ou comprar a partir de uma busca aleatória sem nenhum custo. Dessa forma, o número de firmas não irá interferir no custo

de procura, diferentemente da busca sequencial, quando após observar o mercado sem nenhum custo o consumidor decide por um custo s de procura considerando buscas repetitivas.

Adicionalmente, mantendo a intensidade de procura constante, havendo o aumento do custo de produção, a diferença entre os preços de monopólio e o valor mínimo rentável de p^* diminui e as firmas colocam preços maiores, mas menos dispersos (no caso extremo $c = p^* = v$).

Apesar de não considerarem que os preços se dispersam entre os consumidores e ao longo do tempo, o modelo de análise estática comparativa proposta por Chandra e Tappata (2011) ilustra a importância da intensidade de procura para a dispersão de preços em uma busca não sequencial.

Para o estudo empírico da dispersão, os modelos teóricos, apesar de oferecerem um referencial para definições do comportamento dos preços, não funcionam para todos os mercados. A utilização de uma metodologia empírica é necessária para aprofundarmos o estudo no mercado brasileiro.

A próxima seção irá tratar especificamente das metodologias utilizadas nos estudos empíricos de comportamento de preços. Além de trabalhos que tratam especificamente de dispersão e custo de procura utilizando especificamente o mercado de varejo de combustíveis.

2.3 Estudos empíricos de dispersão de preços

Os estudos empíricos tentam evidenciar em que proporção os custos de informação contribuem para a dispersão de preços. As técnicas e disponibilidades de séries de preços e demais informações de estruturas de mercado são razoavelmente recentes, o que explica porque essa área ainda é pouco explorada na teoria econômica.

Outro ponto de dificuldade é definir qual a área de mercado significativa para a procura do consumidor, logo, qual o mercado de competição relevante para a dispersão. Davis (2005), que trata de preocupações sobre concentrações horizontais, utilizando o preço dos cinemas nos Estados Unidos, criou um modelo empírico para analisar os efeitos de concorrência sobre os preços dos cinemas. Mesmo não sendo um estudo específico sobre dispersão a metodologia econométrica, apresenta uma base para a leitura do efeito de estruturas de mercado sobre os preços em diferentes raios geográficos de competição. Assim, é possível definir quais as áreas de competição entre os postos.

Conforme Davis (2005), a avaliação das estruturas de mercado nos preços, tem uma vasta história dentro da economia industrial, mas evidências econométricas, no entanto, são pequenas. O trabalho utiliza uma rica base de dados do mercado de cinemas norte-americanos e os resultados apontam que o preço de entrada dos filmes depende diretamente da competição local de cinemas. A magnitude dos efeitos estimados, entretanto, é pequena, e não há evidência que grandes redes possam com decisões independentes levar a preços maiores.

Para determinar o tamanho dos mercados dos cinemas, ou seja, a área de influência de um estabelecimento. O autor utiliza raios de vizinhança predeterminados e então é contabilizado o número de estabelecimentos que possuem uma variedade de distâncias de proximidade ao cinema.

Dada a localização de cada estabelecimento $Localização_{mt} = \{L_{gmt}\}_{g=1}^{G_{mt}}$, é calculado o número de lugares com determinada distância de h, sendo m o índice que representa o mercado, t tempo e $g = 1 \dots G$ determinando vizinhos. Definindo, então, $\Gamma_{hmt}(a, b; Localização_{mt}) \equiv \{g \ / \ a \leq d(L_h, L_g) < b\}$ sendo o conjunto de cinemas com distância maior que a e menor que b do lugar h. Assim, se constrói a seguinte equação:

$$Cinema_{hmt}(a, b) = \sum_{\Gamma_{hmt}(a,b)} Vizinhos_{gmt} \tag{03}$$

Dividindo o número de estabelecimentos da mesma empresa no raio de mercado (rede), e o número de estabelecimentos de empresas diferentes (rivais), o modelo tenta atingir seu objetivo de análise de integração horizontal de empresas.

A partir da construção dos mercados com respectivas distâncias, é então definida a estimação utilizando os vetores: $rede_{hmt} = rede_{hmt}(0; 0{,}5); rede_{hmt}(0{,}5; 1{,}0) \dots rede_{hmt}(25; 30))$ e $rival_{hmt} = rival_{hmt}(0; 0{,}5); rival_{hmt}(0{,}5; 1{,}0) \dots rival_{hmt}(25; 30))$. Dessa forma, definida então a equação que irá estimar a variação no preço.

$$p_{hmt} = \beta_0 + \alpha_1 rede_{hmt} + \alpha_2 rival_{hmt} + x_{hmt} + \mu_h \beta + \epsilon_{hmt} \tag{04}$$

A variável x é um vetor de controle, e as variáveis $rede$ e $rival$ são o somatório de postos em determinado raio geográfico, pertencentes ou a mesma rede, ou a uma rede rival.

Dessa forma é estimado o efeito marginal sobre o preço em diferentes raios utilizando o método de painel.

Os resultados permitem analisar se o aumento de empresas rivais, ou empresas do mesmo setor, tem um papel significativo sobre o comportamento dos preços. Além de permitir a comparação entre os efeitos de mais uma firma rival, ou de mais uma firma da rede sobre o preço final cobrado ao consumidor.

Definidos os mercados de relevância competitiva, resta então abordar os estudos que efetivamente mensuram a dispersão de preços. Stigler (1961) foi o primeiro a lançar *insights* sobre o comportamento empírico da dispersão. Para ele, a dispersão se configura como uma função do montante médio de procura, que por sua vez está relacionado com a natureza do produto. O autor define alguns pressupostos básicos e intuitivos sobre o comportamento dos preços: quanto maior o gasto em um produto, maior a economia com a busca por melhores preços, e maior o montante de procura; quanto mais experiência os compradores adquirem com as compras, mais efetiva se torna a procura; quanto maior a fração de compradores que compram repetitivamente, maior a correlação entre preços sucessivos, e maior o montante acumulado de busca; e, por último, o custo de procura é maior quanto maior for a dimensão geográfica.

McAfee (1983) também apresenta em sua conclusão suposições sobre a distribuição dos custos de busca e sobre a natureza das firmas. Sendo que muitas das suposições podem ser investigadas empiricamente, e auxiliam nas previsões intuitivas sobre os mercados.

Supõe-se que o número de firmas seja influenciado pela distribuição das funções de custo de potenciais firmas, e, *ceteris paribus*, haverá mais firmas se maiores forem os intervalos de custos de busca dos consumidores. Também é exposto que, *ceteris paribus*, a variância dos preços é crescente dado um aumento no número de firmas, uma diminuição nas curvas das funções de custo marginal, e por uma redução na função de densidade da distribuição dos custos de procura.

Sobre os modelos empíricos, a utilização de um coeficiente de variação para comparar os preços de forma temporal ou estática, é uma das primeiras abordagens que vislumbram comprovação empírica de dispersão.

Conforme Macminn (1980), se as firmas possuírem uma distribuição marginal uniforme dos custos, então a variância dos preços é dada pela equação:

$$\sigma_p^2 = \left(\frac{n^*-1}{n^*}\right)^2 \frac{(\bar{m}-\underline{m})^2}{12} \qquad (05)$$

Sendo n o número ótimo de firmas e m o custo marginal de cada empresa. A partir da variância chega-se ao coeficiente de variação dado por $CV = \sigma_p / E[p]$. Assim, quanto maior a variância, maior será a dispersão no mercado.

Outra metodologia típica utilizada é o intervalo de preços. Conforme Baye et al. (2006), o *gap* entre os dois menores preços. A intuição sobre esse modelo permite perceber o montante de dinheiro economizado entre comprar no melhor preço, ao invés de fazer uma busca aleatória.

Com relação a modelos de dispersão temporal, o modelo empírico de Chandra e Tappata (2011) observa as mudanças no ranking de preços dos postos de combustíveis americanos. A técnica de *rank reverso* procura evidências de estratégias mistas no mercado de combustíveis. Para isso, a ideia é encontrar que a proporção de tempo que um vendedor mantém a estratégia de cobrar o preço mais alto em períodos consecutivos é baixa. Espera-se que as firmas randomizem suas posições com relação aos preços ao longo do tempo, assim chegariam em uma frequência ótima dadas as escolhas dos competidores.

Para testar o fator da informação imperfeita os autores ressaltam que é necessário um grupo de controle. Em função da prática de exposição dos preços dos combustíveis em frente aos postos, o grupo de controle é formado por estabelecimentos vizinhos, assim os consumidores têm acesso aos preços com custo zero. É então feita a comparação entre a dispersão de preços no mercado que tem custo de procura com o grupo de controle com custo zero.

Espera-se que os preços de dois postos localizados na mesma esquina mostrem somente as diferenças do produto. Ou seja, não há custo de procura na decisão do consumidor por um ou outro posto. Nesse caso, se espera que o *rank* reverso seja menos frequente para estabelecimentos próximos entre si, do que estabelecimentos mais distantes.

Para isso é utilizada uma proporção em que dois postos (*i*,*j*) sobre o tempo T_{ij} tal que seja $pit \geq pjt$ observado em um longo período, é definindo o *rank reverso* entre os postos i e j como a proporção observada em que $pjt > pit$ conforme a equação:

$$r_{it} = \frac{1}{T_{ij}} \sum_{t=1}^{Tij} I_{(pjt > pit)} \qquad (06)$$

Nesse modelo os estabelecimentos são divididos conforme mercados e preços praticados por "concorrentes próximos", estes constituídos por um determinado raio de proximidade.

Os resultados obtidos indicam que combustíveis associados com maiores custos de procura apresentam dispersão de preços em pontos de equilíbrio maiores, ou com maior dispersão. O artigo também conclui que mercados atomizados, ou com maior número de firmas estão associados com uma maior dispersão. Finalmente eles expõem que os consumidores deveriam procurar menos quando os preços sofrem choque e aumentam rapidamente.

Adicionalmente os dados mostraram que mais de 90% dos pares de postos tem um *rank reverso* positivo com média de 0,15, o que significa que os estabelecimentos que normalmente cobram os maiores preços, tiveram o preço mais baixo 15% do tempo. Na estimação o indicador nunca pode ser maior que 0,5, já que a medida de "longo período" considera preços $pit \geq pjt$em pelos menos 50% do tempo estimado.

A partir do *rank* reverso e o grupo de controle, é possível estabelecer uma relação entre a teoria dos modelos de busca e a dispersão observada nos mercados. Os resultados dos autores indicam que a busca do consumidor tem um papel importante na variação temporal de preços.

Marvel (1976), assim como Chandra e Tappata (2011), também utilizou dados dos postos de combustíveis em seu modelo empírico. O estudo, no entanto, buscava preencher um *gap* teórico ao demonstrar a importância da informação nos mercados de varejo de combustíveis. A informação teria impacto em dois aspectos principais do comportamento dos preços: 1) o montante oferecido de dispersão que é observado entre os estabelecimentos dentro de determinados mercados; 2) e a extensão da variação de preços ao longo do tempo.

Para a estimação da variação entre as cidades, Marvel (1976) utilizou a seguinte regressão:

$$\ln r_i = \beta_0 + \beta_1 \ln b_i + \beta_2 \ln y_i + \beta_3 \ln s_i + \beta_4 \ln v_i + \beta_5 \ln m_i + \beta_6 \ln t_i + \delta_i \qquad (7)$$

Em que r_i é o range médio de preços em cada mercado; b_ié o consumo de gasolina por carro; y_i é a renda média das famílias; s_i é a média de educação das famílias; v_i a medida de variação do preço de gasolina; m_i índice de Herfindahl[9], relativo à estrutura do mercado; t_i fator de correção para a diferença de impostos; e δ_i o termo de erro. O autor justifica que as

[9] Considerando que s_j identifica a proporção o *marketshare* da firma j para determiado Mercado. Então, n firmas atuando nesse Mercado definirão o índice Herfindahl da seguinte forma: $H = \sum_{j=1}^{n} s_j^2$.

variáveis de educação e renda poderiam ser utilizadas como *proxies* para o nível e a qualidade de informação dos mercados.

Assim, Marvel (1976) conclui que é aceitável que o intervalo de preços seja menor quanto menor fosse a renda dos compradores, maior o nível educacional, maior considerado o volume de compras, e maior a correlação de preços sucessivos na distribuição de preços.

Finalmente, considerando os diferentes modelos apresentados, o próximo capítulo irá abordar os dados de preços e qual a melhor metodologia a partir deles para mensurar a área de mercado relevante e o nível de dispersão.

3. METODOLOGIA

A metodologia do presente estudo é dividida em três partes. Na primeira são descritos os dados utilizados, a segunda consiste na metodologia de localização dos estabelecimentos e como são traçadas as distâncias, e, por final, as estimações para mercado relevante e mensuração da dispersão.

3.1 Base de dados

3.1.1 Levantamento de preços ANP

Os dados dos preços praticados pelos postos foram cedidos pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis – ANP[10], que realiza uma coleta semanal de preços de varejo desde 2001.

A pesquisa é feita em todos os estados da Federação mais o Distrito Federal, totalizando 555 localidades brasileiras selecionadas a partir de critérios econômicos, tais como renda, população, número de postos revendedores e frota de veículos. A coleta é feita por visita pessoal de pesquisadores treinados aos estabelecimentos, e os dados são registrados em planilhas nas quais os dados são criticados, verificados e consistidos eletronicamente por meio de software específico. [11]

Para o estudo do mercado de postos de combustíveis serão analisados os dados referentes à cidade de São Paulo. A decisão pela região dos dados se deu em função da prática de cartéis, que atingem principalmente municípios menores.

De acordo com a Secretária de Defesa Econômica – SDE/MJ, denúncias de prática de associação de cartel dos postos respondem por um terço do total das denúncias recebidas pelo órgão.[12]Conforme o Conselho Administrativo de Defesa Econômica – CADE, entretanto, não há nenhum registro de processo de cartel em andamento no município de São Paulo[13].Além

[10] Os dados atuais estão disponíveis em: www.anp.org.br/preco/

[11] A metodologia do levantamento de preços é apresentada pela ANP em "Metodologia utilizada para realização da pesquisa de preços no âmbito do Levantamento de Preços e de Margens de Comercialização de Combustíveis da ANP. Vigência: 12/08/2010 a 11/08/2012." Disponível em: http://www.anp.gov.br/?dw=41567. Acesso em 08/03/2013.

[12]Fonte: http://portal.mj.gov.br. Acesso em 08/03/13.

[13] De acordo com dados do CADE foram julgados, estão aguardando julgamento e estão em trâmite no período de 2002 a 2011, casos de cartel nas seguintes cidades brasileiras: Florianópolis- SC, Goiânia- GO, Lages- SC, Belo Horizonte- MG, Brasília- DF, Recife- PE, Guaporé- RS, Londrina- PR, Manaus- AM, Santa Maria- RS, Teresina- PI, Bauru- SP, Vitória- ES, João Pessoa- PB e Caxias do Sul- RS.

disso, a própria escala do munícipio é tal que a manutenção de esquemas colusivos é pouco plausível.

O banco de dados cedido pela ANP inclui as seguintes variáveis:

i) razão social do posto;

ii) local;

iii) bandeira;

iv) preço de revenda da gasolina;

v) preço de distribuição da gasolina;

vi) preço de revenda do etanol;

vii) preço de distribuição do etanol.

As variáveis de preço de distribuição da gasolina e preço de distribuição do etanol não serão utilizadas no estudo em função da coleta não ser regular, portanto, muitas observações não apresentam esses valores. A informação "bandeira" caracteriza a marca que o posto adota muitas vezes sendo atreladas a distribuidora de combustíveis, ou "bandeira branca" quando o posto não está vinculado à nenhuma distribuidora.

O banco de dados inicial trata de 260.204 observações divididas entre os anos de 2001 a 2012[14]. São pesquisadas aproximadamente 3.600 razões sociais diferentes entre os períodos.

As observações fornecidas pela ANP, em função dos registros serem feitos manualmente, tiveram correções feitas nos endereços. Essa opção foi tomada no intuito de melhorar a efetividade do georreferenciamento e aproximar as informações fornecidas pelo Instituto Brasileiro de Geografia e Estatística – IBGE do censo demográfico de 2010, aos dados dos postos.

O maior preço médio atingido pela gasolina e etanol ocorreu no ano de 2011, ano em que, segundo o Boletim Anual de Preços do Petróleo, ocorreram aumentos expressivos dos preços do etanol hidratado e da gasolina no primeiro quadrimestre do ano. Essa situação ocorreu em consequência das majorações nos preços praticados pelos produtores, dado o aumento da frota de veículos bicombustíveis. No entanto, após o início da safra da cana-de-açúcar, quando houve recuo dos preços em todos os segmentos da cadeia produtiva, os preços relativos entre os combustíveis favoreceram, em geral, o consumo de gasolina.

No total os dados têm 133 períodos de coleta, que correspondem aos dados mensais desde 2001 até 2012. As coletas não são constantes ao longo dos anos, logo não

[14]Julho de 2001 a Julho de 2012.

necessariamente o posto terá o seu preço coletado nos 133 períodos. Alguns estabelecimentos, por exemplo, só foram coletados em 1 período.

Conforme observado no gráfico 04, o eixo horizontal apresenta o ano e o eixo vertical o número de coletas, é possível notar que no período de 2004 as coletas estão em maior número. Há também picos menores de coleta em 2005, 2006 e 2007.

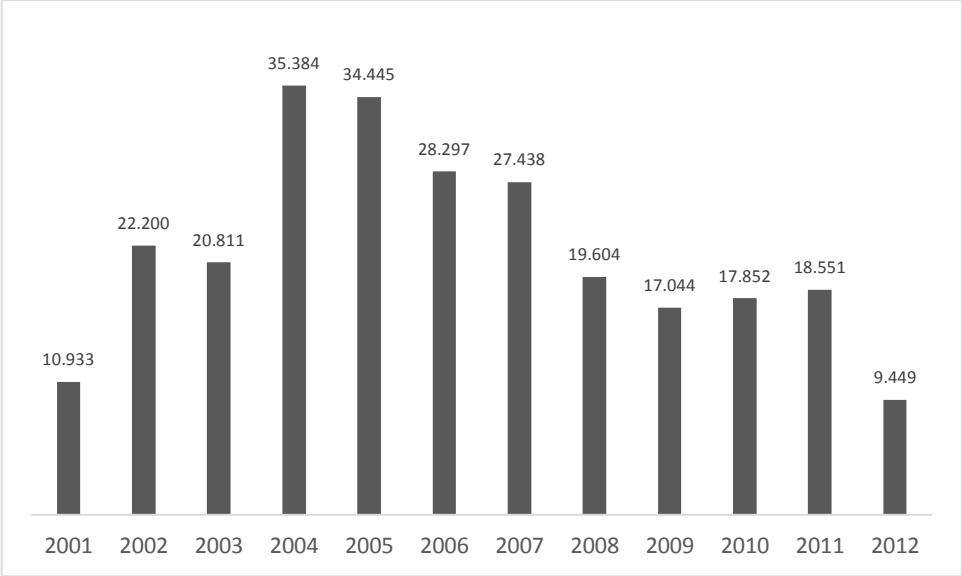


Gráfico 04 - Número de Observações Analisadas por Período. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor

Analisando separadamente como se comporta a dispersão em períodos específicos, nos gráficos 05, 08 e 11, temos no eixo horizontal o preço da gasolina, no eixo vertical o preço do etanol, e cada ponto uma observação coletada. Também é indicado o intervalo de confiança de 95%, uma reta de regressão e a reta pontilhada representando a relação técnica entre gasolina e etanol.

Nos gráficos 06 e 07, 09 e 10, e 12 e 13 são apresentados os histogramas com a dispersão dos combustíveis gasolina e etanol nos mesmos períodos analisados nos gráficos 05, 08 e 11 respectivamente. No eixo horizontal constam os preços, e no eixo vertical a porcentagem de postos que praticou determinado preço nesse período.

Os gráficos 05 e 08 revelam o comportamento dos preços nos períodos com mais postos coletados, e menos postos coletados. Conforme o gráfico 11, é possível ter uma ideia sobre o afastamento dos preços em relação a relação técnica de gasolina e etanol, em período de alta do etanol no mercado brasileiro. Os histogramas reforçam a ideia de dispersão, uma vez que em

nenhum dos períodos foi praticado o mesmo preço por mais de 30% dos postos analisados, seja com relação à gasolina ou etanol.

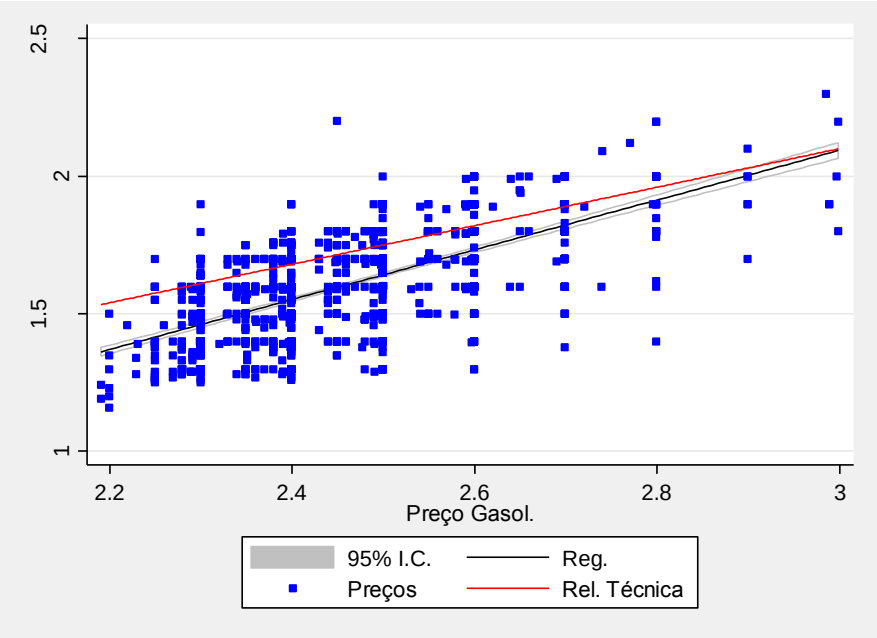
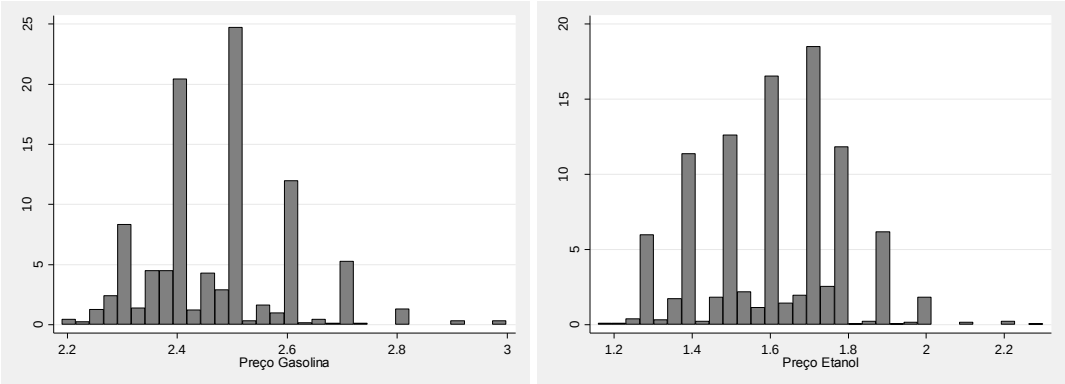


Gráfico 05 - Dispersão de Preços de Venda de Gasolina e Etanol 01/2004. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



Gráficos 06 e 07 – Histograma Dispersão de Preços de Venda de Gasolina e Etanol 01/2004. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor

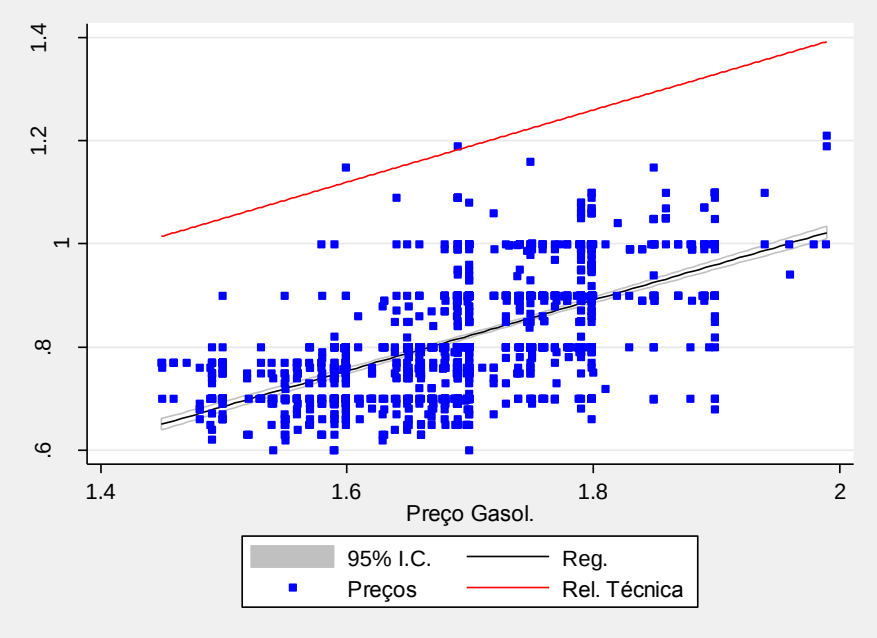
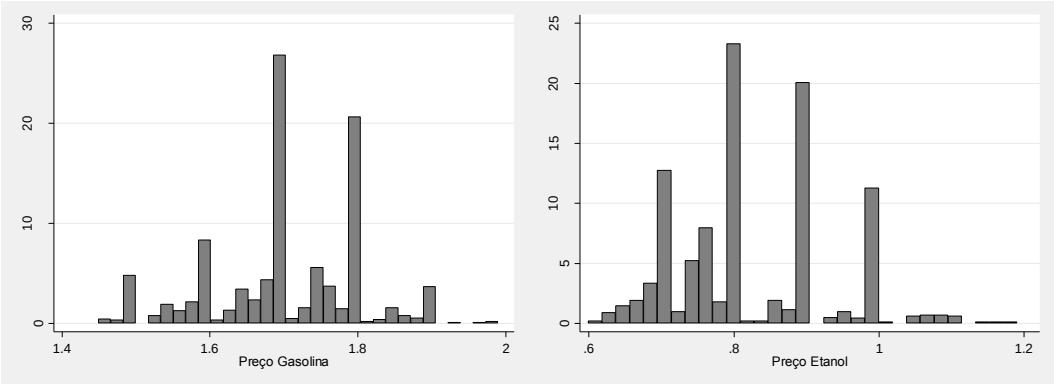


Gráfico 08 - Dispersão de Preços de Venda de Gasolina e Etanol 01/2009. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



Gráficos 09 e 10 – Histograma Dispersão de Preços de Venda de Gasolina e Etanol 01/2009. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor

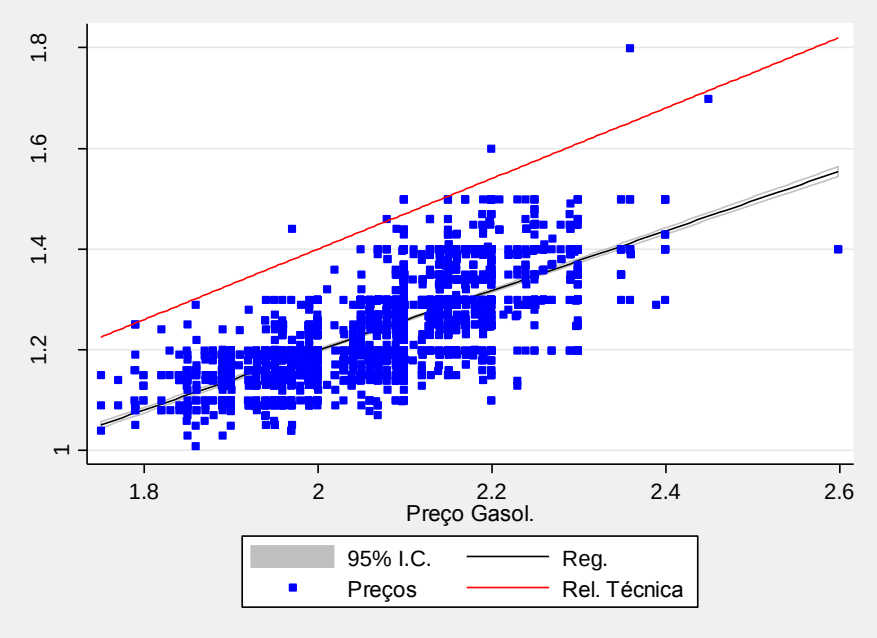
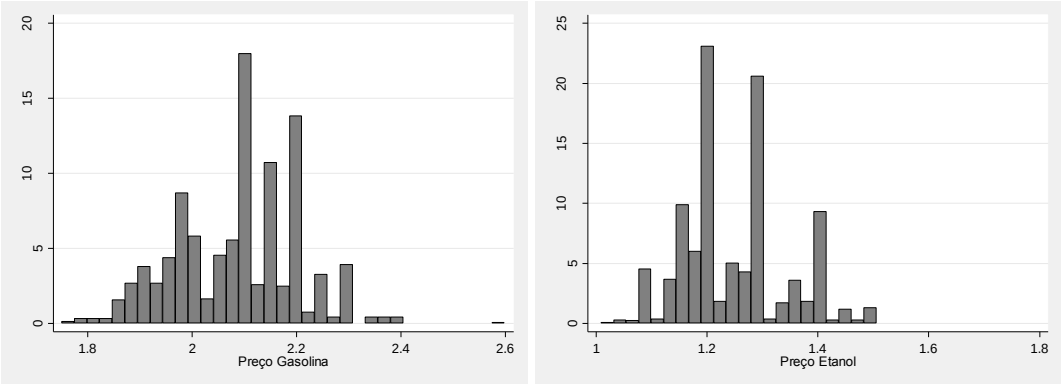


Gráfico 11 - Dispersão de Preços de Venda de Gasolina e Etanol 01/2011. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



Gráficos 12 e 13 - Histograma Dispersão de Preços de Venda de Gasolina e Etanol 01/2011. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor

Apesar de verificarmos que há dispersão entre os estabelecimentos, é necessário examinar qual a fonte desse fenômeno e se a busca do consumidor tem alguma influência. Para isso será necessário saber onde estão localizados os postos. A próxima seção da metodologia trata do georerreferenciamento a partir dos endereços fornecidos pela ANP.

3.1.2 Georreferenciamento

Para as estimações, foi necessário mapear a área do mercado, bem como as distâncias entre cada uma das operações. Para tal, foi utilizado um o sistema de georreferenciamento, que

consiste em tornar suas coordenadas conhecidas em um dado sistema de referência, a partir de uma base de informações geográficas.

De acordo com as informações dos endereços fornecidas pela ANP, foi possível encontrar a latitude e longitude utilizando o sistema de georreferenciamento feito por meio do sistema Google API (ou Google AJAX APIs). Esse sistema em linguagem *JavaScript* desenvolvido pela empresa Google, permite interação entre um sistema da Google e outro software que lê tais dados. Utilizando os endereços (logradouro, número, bairro e cidade), a base de dados do Google Mapas, por meio do programa *Google Maps JavaScript* API v3, fez a leitura das coordenadas e utilizando o pacote de *software Stata* como *callback*, processou o resultado obtido por meio do API.

O uso do sistema de georreferenciamento utilizando o *Google Maps* foi descrito por Ozimek e Miles (2011), pelo comando *geocode* no *Stata*. A utilização do comando, entretanto, não foi possível devido a mudanças da Google Maps. Por essa razão, foi usado o comando *insheetjson* também no *Stata* para a leitura dos dados em *JavaScript*.

O Google API foi utilizado para se fazer o chamado georreferenciamento reverso, que é a tradução de um endereço legível em uma localização em um mapa (coordenadas geográficas). As coordenadas resultantes são retornadas na ordem de melhor para pior aproximação.

Geralmente, o endereço mais exato é o primeiro. Quanto mais detalhado o endereço, fornecendo informações de bairros, cidades, municípios, estados, etc., maior o nível de aproximação da coordenada. Ressalta-se, que uma vez que os endereços são preenchidos manualmente na coleta da ANP, possíveis erros de digitação não detectados no tratamento dos dados podem gerar erros na localização dos postos. O sistema lê todos os endereços e determina coordenadas aproximadas, portanto, caso a localização seja inexistente, foi da mesma forma atribuída uma coordenada.

Com base nos dados de localização foram então mapeados os raios de vizinhança entre os estabelecimentos de combustíveis. Para isso foi utilizado o comando *distmatch* no software *Stata*, que a partir das informações das coordenadas, traçou as distâncias através da formula de *Haversine*[15] que cada posto tem do seu concorrente, dentro de um raio limite preestabelecido.

3.2 Modelo empírico

[15] Equação normalmente usada para navegação, a distância de uma esfera é calculada a partir dos pontos das coordenadas. Fórmula utilizada no comando *distmatch*.

No modelo empírico, consideramos que para efetuar a compra de combustível, o consumidor usa como informação para o melhor preço a amostra das observações feitas em seu trajeto diário. Como é obrigatória nos postos a exposição dos preços, é plausível que cada consumidor ao seguir seu trajeto os observe e compare. Os consumidores que não seguem esse comportamento não possuem informações necessárias para efetuar a compra no estabelecimento com menor preço, e irão adquirir o combustível de forma aleatória.

A análise empírica é dividida em duas etapas, correspondentes a duas metodologias distintas. A primeira é baseada nas estimações de Davis (2006) para determinarmos a extensão geográfica dos mercados relevantes. Já a segunda etapa é baseada no trabalho de Chandra e Tappata (2011).

Na tabela 01 estão divididos os raios de distâncias que serão analisados, utilizando como referência os raios de competição de Davis (2006)[16]:

Raio 01	0,0 a 1,0 km
Raio 02	1,0 a 2,0km
Raio 03	2,0 a 3,0km
Raio 04	3,0 a 4,0 km
Raio 05	5,0 a 6,0 km
Raio 06	6,0 a 7,0 km
Raio 07	7,0 a 8,0 km
Raio 08	8,0 a 9,0 km
Raio 10	9,0 a 10,0 km

Tabela 01–Raios de competição. Elaboração: Autor

Dentro do raio 01 estarão os estabelecimentos vizinhos com até 1,0km de distância entre eles, no raio 02 estabelecimentos com 1,0 km a 2,0 km de distância e assim sucessivamente. Espera-se verificar se o número de estabelecimentos no entorno, tem significância sobre o preço cobrado nos diferentes raios. Em outras palavras, queremos verificar se existe um ambiente competitivo importante nos raios analisados.

Davis (2006), é a base da metodologia para estabelecer os efeitos do número de estabelecimentos sobre os preços, mas diferentemente do modelo do autor, são acrescentadas

[16] Davis (2005) utiliza raios com intervalos de 1km até o raio 10 (9km a 10km), ampliando para intervalos de 5km em raios superiores.

dummies que identificam a marca do estabelecimento, tempo e bairro, e não são divididos os postos entre rivais e postos da mesma rede. São também incluídos, como variáveis de controle, dados do censo 2010 relativos à renda e tamanho da população dos bairros. O controle por essas variáveis vai melhorar a mensuração do efeito da competição sobre os preços, ou seja, o efeito do número marginal de concorrentes nas vizinhanças.

A partir dos raios de vizinhança é feito um somatório com o número de vizinhos de cada estabelecimento i em cada raio r por período de tempo t [17].

Para todos os postos então, teremos uma vizinhança tal que: $Vizinhos_{irt} = Vizinhos_{i1t}(0;1{,}0); Vizinhos_{i2t}(1{,}0;2{,}0) \ldots Vizinhos_{i10t}(9{,}0;10{,}0)$. Montados os raios de vizinhança é estimada a seguinte regressão baseada em Davis (2005):

$$p_{it} = \beta_0 + \beta_r Vizinhos_{irt} + \beta_2 X + \varphi_i + \varepsilon_{it} \qquad (08)$$

Sendo p_{it} as variáveis de preços, β_0 o intercepto, $Vizinhos_{irt}$ a variável relativa ao somatório do número de vizinhos por raio; X vetor das variáveis de controle: renda média das famílias por bairro; densidade demográfica por bairro, período; e *dummies* de bandeira[18] e bairro; φ_i representa um efeito fixo em relação ao tempo e ε_{it} o termo de erro.

Dada a significância das variáveis nos diferentes raios, a segunda parte da metodologia será determinar o grau de dispersão utilizando como referência a metodologia de *rank reverso* de Chandra e Tappata (2011). Assim o *rank reverso* r_{it} em cada mercado será definido a partir do índice:

$$r_{it} = \frac{1}{T_{ij}} \sum_{t=1}^{Tij} I_{(pjt>pit)} \qquad (06)$$

Em que os preços entre dois postos (*i,j)* localizados no mesmo raio, sobre o tempo T_{ij}tal que usualmente[19]$pit \geq pjt$. Definindo o *rank reverso* entre os postos ie $j$como a proporção observada em que $pjt > pit$conforme equação (06). Assim o índice revela, dentre os postos que normalmente cobram o maior preço, com que frequência irão cobrar o menor preço.

[17]Os dados fornecidos pela ANP, não fornecem uma amostra constante. As informações dos preços dos postos, não necessariamente, são coletadas em todos os períodos, assim o número de vizinhos se altera no tempo.
[18]As *dummies* de bandeira conforme anexo 01.
[19]O tempo definido em que $pit \geq pj$deverá ser no mínimo 50% do período observado. O rank reverso, portanto, não poderá superar 0,5.

O *rank reverso* é avaliado a partir de duas condições: nos raios em que os vizinhos têm significância sobre os preços, o que é definido na metodologia de Davis (2005) conforme equação (08); e nos estabelecimentos de "esquina" ou amostra de controle, onde consideramos uma distância mínima de 100m.

A amostra de controle nos permite avaliar uma condição onde o custo de informação, ou o custo de procura, é zero. Considerando que os preços estão expostos na frente dos estabelecimentos, a informação dos consumidores é completa ao decidirem pela opção mais vantajosa.

Já nos raios selecionados a partir da equação (08), será possível verificar qual o nível de dispersão, e analisar se a teoria e os modelos de preços estão adequados aos resultados empíricos desse mercado.

4. RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados das duas etapas da análise, sendo a primeira baseada no artigo de Davis (2005), a fim de verificar os raios de competição do mercado de varejo de combustíveis. A segunda etapa consiste na análise do *rank reverso*, conforme Chandra e Tappata (2011), utilizando os raios de competição estimados na primeira etapa.

4.1 Raios de competição

A estimação da primeira parte da metodologia foi feita com base na equação (08), utilizando modelo em painel de mínimos quadrados ordinários, efeitos fixos[20]e variáveis instrumentais.

$$p_{it} = \beta_0 + \beta_r Vizinhos_{irt} + \beta_2 X + \varphi_i + \varepsilon_{it} \qquad (08)$$

As tabelas 02 e 03 apresentam os valores dos coeficientes β_r das variáveis $Vizinhos_{irt}$, sendo $r = (1 \dots 10)$.

Para as estimações em tabela 02 se assume que as estruturas de mercado, representadas pelo número de vizinhos, não estejam correlacionadas com o erro, utilizando a metodologia de mínimos quadrados ordinários e efeitos fixos. Os resultados da tabela 03 utilizam variáveis instrumentais, flexibilizando a condição utilizada anteriormente.

Os resultados da primeira etapa consideraram somente a primeira observação da razão social por mês, sendo os preços de etanol e gasolina a média dos valores no período. Assim, como a base não é censitária, ou seja, não necessariamente todos os postos são pesquisados em todos os períodos, foram utilizados os valores médios e somente uma observação por mês.

As colunas (1.1) e (1.2) da tabela 02, representam os resultados utilizando o método de mínimos quadrados ordinários para gasolina e etanol respectivamente. Já as colunas (2.1) e (2.2) apresentam as estimações de efeito fixo.

Os raios 04 e 01 apresentam resultados estatisticamente sensíveis ao número de vizinhos para gasolina, utilizando-se o método de mínimos quadrados e efeitos fixos respectivamente. Com relação ao etanol, as estimações por mínimos quadrados, com exceção do raio 04, são

[20] Considera-se que efeitos individuais sejam correlacionados com os regressores, não justificando a adoção de estimações por efeitos aleatórios.

significativas com até 10% de significância. Já nas estimações com efeito fixo os raios 01, 02, 05, 07 e 10 são significativos a 5%. Nota-se também que os resultados mostram coeficientes positivos nos primeiros raios. A magnitude do efeito é baixa nas estimações, não ultrapassando R$ 0,01 por litro de combustível, conforme o valor do raio 01 para etanol.

	MQO (1.1)	**MQO (1.2)**	**Efeitos Fixos (2.1)**	**Efeitos Fixos (2.2)**
	Gasolina	**Etanol**	**Gasolina**	**Etanol**
∑ Vizinhos (Raio 01)	0,002225	0,007974***	0,004622***	0,010911***
	(1,575)	(5,617)	(3,064)	(7,472)
∑ Vizinhos (Raio 02)	0,000580	0,002518***	0,001259	0,002872***
	(0,761)	(2,987)	(1,605)	(3,723)
∑ Vizinhos (Raio 03)	0,000690	0,001135*	0,000637	0,000877
	(1,075)	(1,730)	(0,972)	(1,404)
∑ Vizinhos (Raio 04)	0,000880*	0,000251	0,000016	-0,000572
	(1,822)	(0,522)	(0,032)	(-1,197)
∑ Vizinhos (Raio 05)	-0,000016	-0,000856*	-0,000346	-0,001088***
	(-0,036)	(-1,895)	(-0,840)	(-2,864)
∑ Vizinhos (Raio 06)	0,000176	-0,000571*	0,000362	-0,000122
	(0,498)	(-1,667)	(0,961)	(-0,346)
∑ Vizinhos (Raio 07)	0,000265	-0,000771**	0,000050	-0,000920***
	(0,740)	(-2,202)	(0,149)	(-2,840)
∑ Vizinhos (Raio 08)	-0,000133	-0,000786**	0,000299	-0,000286
	(-0,454)	(-2,526)	(1,011)	(-0,916)
∑ Vizinhos (Raio 09)	-0,000057	-0,000480*	0,000430	-0,000049
	(-0,212)	(-1,674)	(1,626)	(-0,175)
∑ Vizinhos (Raio 10)	-0,000254	-0,001368***	0,000154	-0,000961***
	(-0,630)	(-3,718)	(0,393)	(-2,653)
R2	**0,115**	**0,175**	**0,124**	**0,168**
Número de Observações	**32.799**	**32.799**	**32.799**	**32.799**

Tabela 02 – Estimações Raios de Proximidade – Mínimos Quadrados Ordinários e Efeitos Fixos. ***Significativo a 1% ** Significativo a 5% *Significativo a 10%. Erro padrão entre parênteses.

Levando em consideração o fato de que regiões mais prósperas tendem a ter mais estabelecimentos e maiores preços, se considera que haja correlação diferente de zero entre o termo de erro e os regressores analisados.

Assim, da mesma forma que em Davis (2005), se opta pela utilização de variáveis instrumentais no intuito de minimizar o problema da endogeneidade. Para a escolha de informações que se tornem bons instrumentos é necessário que a correlação entre o erro e a variável escolhida seja zero, mas que haja correlação entre ele e a variável instrumentalizada.

O uso de estruturas de mercado com defasagem de 4 períodos, assim como no artigo base da metodologia, foi a alternativa escolhida neste trabalho.

É utilizado o preço médio de gasolina ou etanol dos bairros, com exceção do bairro da observação p_i, ou seja, o valor médio de preço praticado pelos bairros vizinhos. Também foram usados: o número de vizinhos nos raios mais próximos (raio 01 até o raio 04), e as médias de renda e densidade demográfica dos bairros vizinhos.

Conforme anteriormente citado, postos com mais vizinhos poderão vir a ter preços mais altos, pois estão estrategicamente melhor posicionados. Assim, as variáveis que representam estruturas de mercado, intentam em demonstrar o mercado onde os estabelecimentos estão inseridos.

As estimações nas colunas (3.1) e (3.2) da tabela 03 mostram o uso de variáveis instrumentais utilizando a quarta defasagem das estruturas de mercado. Cada linha da tabela representa um coeficiente β_r na equação (08).

Podemos analisar que os resultados do raio 04 até o raio 07 são significativos, apresentando magnitude negativa nos raios 05 e 07. Os valores relativos aos testes de Sargan-Hansen[21] e o teste de significância conjunta das variáveis, entretanto, rejeitam a hipótese nula, mostrando que o conjunto de coeficientes não é significativo.

No intuito de restringir o mercado para a análise da dispersão, se optou por utilizar a área até o raio 07.

	Variáveis Instrumentais (3.1)	**Variáveis Instrumentais (3.2)**
	Gasolina	**Etanol**
∑ Vizinhos (Raio 01)	-0,169791	-0,205862
	(-1,140)	(-1,287)
∑ Vizinhos (Raio 02)	0,044549	0,042391
	(0,636)	(0,564)
∑ Vizinhos (Raio 03)	0,024400	0,057069
	(0,298)	(0,631)
∑ Vizinhos (Raio 04)	0,274669*	0,296874*
	(1,741)	(1,734)
∑ Vizinhos (Raio 05)	-0,444784**	-0,530444**
	(-2,163)	(-2,188)
∑ Vizinhos (Raio 06)	0,801099***	0,954054**
	(2,638)	(2,556)

[21] No teste de Sargan-Hansen a hipótese nula conjunta é que os instrumentos são instrumentos válidos, ou seja, não correlacionados com o termo de erro, e que os instrumentos excluídos são corretamente. A rejeição põe em dúvida a validade dos instrumentos.

∑ Vizinhos (Raio 07)	-0,811543**	-0,808950**
	(-2,513)	(-2,283)
∑ Vizinhos (Raio 08)	0,096471	-0,153851
	(0,288)	(-0,390)
∑ Vizinhos (Raio 09)	-0,032443	0,103392
	(-0,110)	(0,303)
∑ Vizinhos (Raio 10)	0,224197	0,222918
	(1,064)	(1,009)
R2	**-0,001**	**-0,001**
Número de Observações	**32.751**	**32.751**

Tabela 03 – Estimações Raios de Proximidade com Variáveis Instrumentais. ***Significativo a 1% ** Significativo a 5% *Significativo a 10% Erro padrão entre parênteses.

4.2 Dispersão temporal de preços

Na segunda etapa da metodologia é feita a estimação do modelo com variável dependente, sendo o *rank reverso* para os raios significativos definidos na etapa anterior.

Primeiramente são organizados todos os pares de estabelecimentos concorrentes (*i,j)* nos raios 01 a 07 e em cada observação de tempo T_{ij}. Assim é possível comparar como os estabelecimentos que são vizinhos se comportam ao longo do tempo com relação a seus concorrentes observados.

Consideramos pit o preço praticado pelo estabelecimento i no tempo T_{ij}, e pjt o preço praticado pelo estabelecimento $j$no tempo T_{ij}.

Após definidos os pares é calculado em quantos períodos T_{ij}, em relação ao total de períodos observados entre (*i,j)*,o posto *i* se comporte tal que $pit \geq pjt$, ou seja, qual a proporção de tempo em que ele irá cobrar um preço maior do que o seu concorrente. Se essa proporção é habitual[22], calcula-se então o *rank reverso* conforme equação (06), ou seja, a proporção no tempo em que $pjt \geq pit$.

Dessa forma o *rank reverso* será proporção de tempo em que o posto, que normalmente pratica o preço maior, irá ter o menor preço. Em outras palavras, dentre os postos que normalmente cobram o maior preço, com que frequência irão mudar de posição cobrando o menor preço[23]

[22]Mais de 50% do tempo.
[23]As estimações consideraram apenas pares de postos que possuíam uma frequência de cinco ou mais períodos observados.

Os gráficos 08 e 09 apresentam dois histogramas, o primeiro com os resultados considerando a proporção com *rank reverso* igual a zero, e o segundo considerando somente os resultados diferentes de zero para o raio 07 (raios de 01 a 06 - vide apêndice 1).

Pode-se observar que pelo menos mais da metade das observações, possuem *rank reverso* positivo em todos os raios, ou seja, estabelecimentos que usualmente cobram o maior valor em algum momento apresentam preços menores do que os seus vizinhos. A proporção de resultados positivos é sempre menor no mercado de etanol, e podemos observar que ela cresce com o aumento dos raios, ou seja, a dispersão parece ser maior no mercado de etanol, e diminuir com as distâncias.

Em razão das observações não serem coletadas de forma censitária, o *rank reverso* apresenta regiões acumuladas em todos os raios observados. Fato que é diferente nas observações do estudo de Chandra e Tappata (2011), onde é utilizada uma base censitária dos preços.

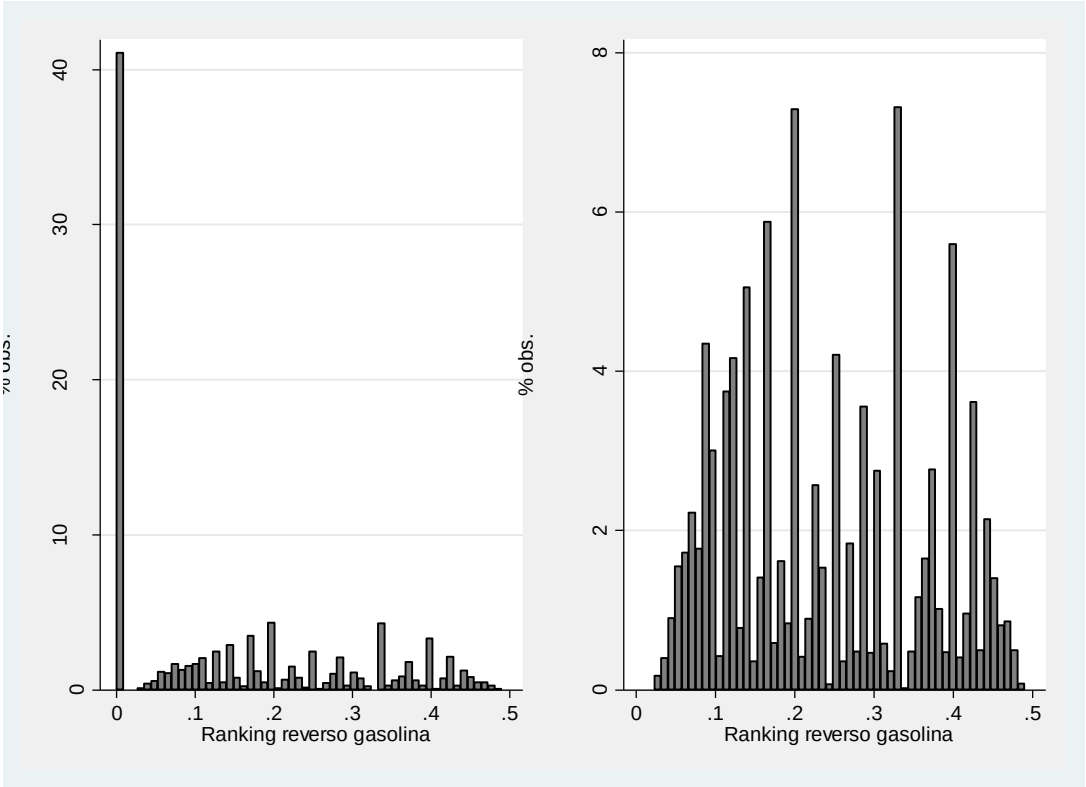


Gráfico 14 – *Ranking reverso* gasolina (6,0 a 7,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor

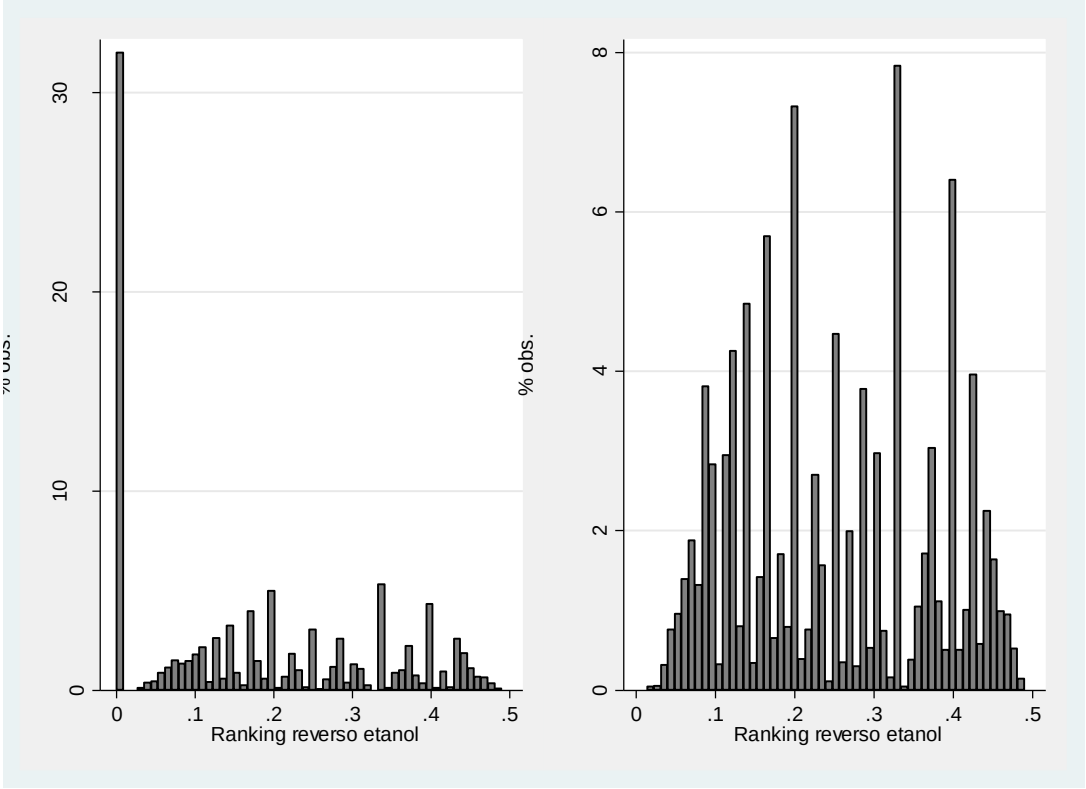


Gráfico 15 – *Ranking reverso* etanol (6,0 a 7,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor

Na tabela 04 temos o resumo do *rank reverso* médio de cada raio e a diferença entre os preços de origem e destino. A tabela também apresenta os resultados para os postos que estejam localizados com maior proximidade[24] classificados como "Esquina".

Os estabelecimentos de "esquina" serão utilizados como amostra de controle para a busca do consumidor. Considerando que os preços estão expostos na frente dos estabelecimentos, os consumidores terão acesso visual aos preços e poderão fazer a sua escolha pelo menor preço.

Os resultados da tabela 04 nos mostram que as observações em que a assimetria de informação é controlada apresentam maior dispersão. Resultado que contraria os modelos de dispersão por busca do consumidor, onde quanto menos clientes informados maior a dispersão de preços.

Entre os raios 01 a 07 também se pode notar que há aumento do *rank reverso*, se estabilizando a partir do raio 04.

Nota-se também que a diferença nos preços é menor nos postos da amostra de controle "Esquina". Assim, mesmo sendo o *rank reverso* maior, o *gap* médio de preços entre os pares é

[24]Consideramos distância igual ou inferior a 100m.

menor, ou seja, postos mais próximos tem maior dispersão, mas a diferença entre os preços é menor.

		Número de Pares	*Rank Reverso* Médio	Diferença Média de Preço (preço de origem - preço de destino)
"Esquina"	**Gasolina**	**3.575**	**0,178**	**-0,084**
	Etanol	**3.854**	**0,220**	**-0,092**
Raio 01	Gasolina	160.926	0,168	-0,099
	Etanol	156.520	0,190	-0,107
Raio 02	Gasolina	374.353	0,157	-0,105
	Etanol	350.672	0,186	-0,109
Raio 03	Gasolina	435.842	0,152	-0,106
	Etanol	421.047	0,177	-0,111
Raio 04	Gasolina	459.705	0,146	-0,108
	Etanol	445.947	0,176	-0,111
Raio 05	Gasolina	468.885	0,145	-0,107
	Etanol	452.639	0,175	-0,111
Raio 06	Gasolina	499.126	0,143	-0,109
	Etanol	481.929	0,173	-0,112
Raio 07	Gasolina	490.208	0,141	-0,111
	Etanol	474.716	0,170	-0,113

Tabela 04 –Resultados *rank* reverso. Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor

Para uma verificação mais robusta da influência da busca do consumidor, assumimos que os postos localizados com distância inferior a 0,1 km apresentam a distribuição de *rank reverso* $F_1(r)$ e postos com distâncias superiores, $F_0(r)$. Dessa forma, se a busca do consumidor tem influência então $F_0(r) < F_1(r)$, ou seja, a dispersão de preços conforme a teoria deve ser menor em estabelecimentos que não tenham assimetria de informação, ou seja, estabelecimentos vizinhos.

Utilizando o teste Kolmorogov-Smirov (K-S) onde a hipótese nula $H_0: F_0(r) = F_1(r)$, e hipótese alternativa $H_1: F_0(r) \cong F_1(r)$, podemos verificar se existe diferença entre as distribuições de *rank reverso* dos estabelecimentos vizinhos e não-vizinhos.

Para reforçar a existência de diferença no *rank reverso* entre estabelecimentos da amostra de controle temos os resultados da equação:

$$r_{it} = \beta_0 + \beta_1 I_{it} + \beta_2 X_{it} + \varepsilon_{it} \quad (09)$$

Onde I_{it} é uma *dummy* para postos com distância inferior a 0,1 km, e X_{it} contém variáveis de controle renda e densidade de habitantes por bairro, conforme censo 2010. Espera-

se que os estabelecimentos apresentem resultado negativo e significativo para I_{it}, assim, o fato dos estabelecimentos serem vizinhos influenciaria negativamente o coeficiente r_{it} de *rank reverso*.

Os resultados conforme o quadro 01, associados aos valores de *rank reverso* (tabela 04), no entanto, nos mostram que os postos de controle, ou os postos "de esquina", apresentam um coeficiente de *rank reverso* maior do que os postos não vizinhos.

	Raio < 1,0 km			
				p-value
Gasolina	**Dummy I_{ij}**		0,009	0.001
		0	0,063	0,000
		1	-0,018	0,094
	(K-S)		**0,063**	**0,000**
Etanol	**Dummy I_{ij}**		0,028	0,000
		0	0,099	0,999
		1	0,000	0,000
	(K-S)		**0,099**	**0,000**

Quadro 01 – Resultados estimações *rank* reverso. "Esquina" *versus* Mercado. Elaboração: Autor. Nota: $H_\alpha: F_0(r_{ik}) = F_1(r_{ij})$ onde $d_{ik} < 0{,}1km$ e $0{,}1km < d_{ij} < 1{,}0km$; $H_\alpha(0): F_0(r_{ik}) > F_1(r_{ij})$; $H_\alpha(1): F_0(r_{ik}) < F_1(r_{ij})$; $H_\alpha(KS): F_0(r_{ik}) \cong F_1(r_{ij})$.

Esse resultado indica que fatores de choques de oferta podem ter uma influência maior do que a assimetria de informação. Certamente os postos que foram selecionados na amostra com vizinhos a menos de 0,1 km apresentam um nível de competição mais elevado, o que reflete no coeficiente de *rank reverso*. Além disso, outros fatores de heterogeneidade entre os estabelecimentos, como a confiabilidade do consumidor, estruturas ou serviços auxiliares, lojas conveniências, ou serviços adicionais, podem também influenciar na dispersão.

Portanto, conforme apresentado, somente a busca do consumidor não explica a dispersão de preços, ou seja, é possível que a diferenciação de produtos, discriminação de preços e choques de oferta tenham um papel mais relevante.

5. CONCLUSÃO

O presente trabalho tratou de identificar e mensurar a dispersão de preços no mercado de postos de gasolina e etanol na cidade de São Paulo. Foram apresentados modelos teóricos de busca do consumidor, e realizados testes empíricos a partir das metodologias de Davis (2005) e Chandra e Tappata (2011).

Para as estimações foram utilizados dados de preços de combustíveis no varejo fornecidos pela ANP, e foi usado o referenciamento geográfico a partir de programa oferecido pelo Google API.

Na primeira parte do trabalho os modelos teóricos de dispersão sugerem que um aumento na quantidade de consumidores informados resultaria em diminuição no nível de preços ótimo. Considerando que mais consumidores tenham acesso a informação, as chances das firmas cobrarem preços maiores e atraírem compradores diminui. Logo, se todos os consumidores são informados, os preços tendem ao custo marginal de produção. Além disso, foi também apresentado um modelo não-monotônico que utiliza a variável intensidade de procura de Chadra e Tappata (2011). Dessa forma dependendo do equilíbrio da proporção de consumidores informados e desinformados, alterações na intensidade de procura teriam efeitos diferentes sobre os preços. Situações intermediárias de intensidade terão preços maiores do que situações de total informação dos consumidores, ou total desinformação.

Na segunda etapa do trabalho, que consiste nos modelos empíricos, foi necessário determinar qual o tamanho das regiões competitivas dos postos de combustível. Dessa forma, a partir da metodologia de Davis (2005), foram estimados os raios de competição.

Os resultados mostraram que até o raio 7 (6,0 km a 7,0 km) teremos um ambiente competitivo em que a existência de mais um estabelecimento tem impacto negativo sobre os preços da gasolina. Para isso foram utilizadas variáveis instrumentais que consideram características médias com quatro defasagens de tempo dos bairros vizinhos aos estabelecimentos.

Utilizando o limite geográfico estimado, com base na metodologia de Davis (2005), foi calculado o *rank reverso* de acordo com Chandra e Tappata (2011). Os resultados de *rank reverso* representam uma tentativa de mensurar a dispersão do mercado, e foram calculados para os raios de 01 (0,0 a 1,0 km) a 07 (6,0 a 7,0 km) e para uma amostra de controle com estabelecimentos vizinhos com menos de 0,1 km de distância.

Para os compradores de combustíveis o raio de busca ótimo não ultrapassa o raio 03 (2,0 km a 3,0 km), já que a dispersão é constante a partir dessa distância. Assim, o motorista que quer pesquisar qual o melhor preço não precisaria percorrer mais que 3,0 km, considerando que nos raios superiores a chance de haver mudanças nos preços é menor. Analisando a velocidade média no horário de pico da tarde em São Paulo de 18,5 km/h[25], o tempo mais eficiente de procura pelo melhor preço não deveria superar cerca de 10 minutos. O *rank reverso* positivo também evidencia a existência de estratégias mistas entre as firmas.

Os resultados obtidos com a amostra de controle, no entanto, não vão de encontro ao esperado na teoria. Para o município de São Paulo, os preços de estabelecimentos próximos tendem a ter maior dispersão, sugerindo que a busca do consumidor não teria grande influência nas estratégias das firmas.

Outro ponto a ser ressaltado é o fato de que entre estabelecimentos vizinhos a possibilidade de dispersão tende a ser maior, mesmo sendo menor a diferença média entre os preços. Deste modo, os consumidores, ao pesquisarem os preços, devem levar em consideração a chance de postos de esquina variarem constantemente sua posição.

Entretanto, se considerarmos os modelos não-monotônicos de busca, é possível que a distribuição de consumidores informados em cada raio possa influenciar os equilíbrios de dispersão e preços. Em função da distribuição ótima de consumidores ser desconhecida nos raios, não é possível compararmos os resultados com a amostra de controle que apresentou maior dispersão e, portanto, maiores ganhos de procura.

Para a análise dos resultados é ainda necessário acrescentar o fato de que os estabelecimentos de venda de varejo de combustíveis oferecem, além do produto principal, conveniências como lojas, serviços gratuitos e programas de fidelização do consumidor. Apesar das estimações tentarem controlar esses fatores através das bandeiras dos postos, as estruturas próprias dos estabelecimentos podem exercer um papel muito forte na estrutura de dispersão de preços e na escolha dos consumidores.

Dessa forma esse estudo conclui que os raios de competição no mercado de varejo de combustíveis se limitam a distância de até 7km entre os estabelecimentos. Já a procura ótima do consumidor, levando em consideração a dispersão, é de em média 10 minutos em horários de pico. Sobre a dispersão e busca do consumidor, apesar de não haver uma conclusão que

[25]Pesquisa de Monitoração da Fluidez. Companhia de Engenharia de Tráfego-CET. Março 2013.

considere a distribuição de consumidores informados, o estudo conclui que a heterogeneidade dos postos tenha um impacto forte sobre os preços, não havendo dispersão significativa relacionada a busca do consumidor.

6. BIBLIOGRAFIA

BOLETIM ANUAL DE PREÇOS 2012. **Preços de Petróleo, Gás Natural e Combustíveis nos Mercados Nacional e Internacional.** Rio de Janeiro: Rio de Janeiro: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP, 2012, 105 p.

BERNHARD, S. 2013**. TRAVELTIME3: Stata command to retrieve travel time and road distance between two locations**. Disponível em: www.fmwww.bc.edu/repec/bocode/t/traveltime3.sthlp. Acessado em 16 dez. 2013.

BAYE, M.; MORGAN, J.; SCHOLTEN, P. Information, Search, and Price Dispersion. In: HENDERSHOTT, T. (ed.). **Handbook of Economics and Information Systems**. Amsterdam: Elsevier Press, 2006.

CARLSON, J. A.; MCAFEE, R. P. Discrete Equilibrium Price Dispersion. **The jornal of political Economy**, Chicago, v. 91, n. 3, p. 480-493, 1983.

CHANDRA, A.; TAPPATA, M. Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Market. **Rand Journal of Economics**, Santa Monica, v. 42, n. 4, p. 681-704, 2011.

COMPANHIA DE ENGENHARIA DE TRÁFEGO-CET, 2013. **Pesquisa de Monitoração da Fluidez. Desempenho do Sistema Viário Principal Volume e Velocidade - 2012 Março/2013**. Disponível em: www.cetsp.com.br/media/228058/2012%20%20volumes%20e%20velocidades.pdf. Acessado em 15 out. 2013.

DAVIS, P. The effect of local competition on retail prices: The U.S. motion picture exhibition market. **Journal of Law and Economics**, Chicago, v. 48, n. 2, p. 677, 2005.

DIAMOND, P. A. A model of price adjustment. **Journal of Economic Theory**, Ithaca, v. 3, n. 2, p. 156-168, 1971.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE, 2010. **Censo Demográfico 2010**. Brasília, 2010. Disponível em: www.censo2010.ibge.gov.br. Acessado em 15 out. 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Base de informações do censo Demográfico 2010: resultados do Universo por setor censitário**. Rio de Janeiro: Documentação do arquivo, 2011, 200 p.

HOSKEN, D. S.; MCMILLAN, R. S.; TAYLOR, C. T. Retail gasoline pricing: What do we know?," **International Journal of Indust**
rial Organization, v. 26, n. 6. p. 1425-1436, 2008.

MARVEL, H. P. The Economics of Information and Retail Gasoline Price Behavior: An Empirical Analysis. **Journal of Political Economy**, Chicago, v. 84, n. 5, p. 1033-1060, 1976.

MACMINN, R. D. "Search and Market Equilibrium". **Journal of Political Economy**, Chicago, v. 88, n. 2, p. 308-327, 1980.
NUNES, C.; GOMES, C. Aspectos concorrenciais do varejo de Combustíveis no Brasil. In: ENCONTRO NACIONAL DE ECONOMIA (ANPEC), 33, 2005, Natal. **Anais**... Natal: ANPEC, 2005. p. 1-19.

PEDRA, D. P.; BICALHO, L. M. N. O.; VILELA, O. A.; BARAN, P. H.; PAIVA, R. M. **Metodologia adotada pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis para detecção de cartéis Coordenadoria de Defesa da Concorrência Agência Nacional do Petróleo, Gás Natural e Biocombustíveis**. Rio de Janeiro: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP, 2010, 20 p.

STIGLER, G. The Economics of Information. **Journal of Political Economy**, Chicago, v. 69, n. 3, p. 213-225, 1961.

SALOP, S.; STIGLITZ, J. E. Bargains and Ripos: A Model of Monopolistically Competitive Price Dispersion. **Review of Economic Studies**, Stockholm, v. 44, n. 3, p. 493-510, 1977.

TIROLE, J. **The Theory of Industrial Organization.** Cambridge: MIT Press, 1988. 473 p.

VARIAN, H. R. A Model of Sales. **American Economic Review**, Pittsburg, v. 70, p. 651-659, 1980.

SHEPARD, A. Price Discrimination and Retail Conguration. **Journal of Political Economy**, Chicago, v. 99, n. 1, p. 30-53,1991.

SHILONY, Y. "Mixed Pricing in Oligopoly". **Journal of Economic Theory**, Ithaca, v. 14, p 373-388, 1977.

STAHL, D. O. Oligopolistic Pricing with Sequential Consumer Search. **American Economic Review**, Pittsburg, v. 79, p. 700-712, 1989.

OZIMEK, A.; MILES, D. Stata utilities for geocoding and generating travel time and travel distance information. **The Stata Journal**, v. 11, n. 1, p. 116-119, 2011.

VARIAN, H. R. Microeconomia: Princípios Básicos – Uma abordagem moderna. 7ª Edição. Rio de Janeiro: Elsevier Editora ltda, 2006, 807 p.

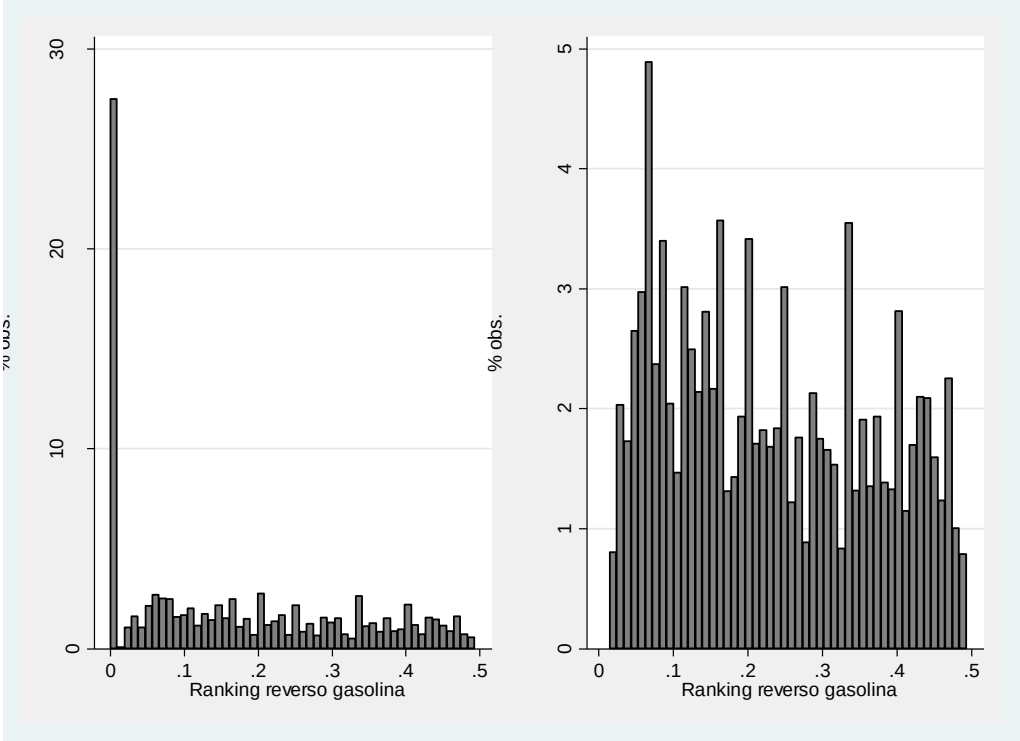
WADA, R. 2009. **Distance matching based on latitudes and longitudes**. Disponível em www.fmwww.bc.edu/repec/bocode/d/distmatch.html. Acessado em 16 dez. 2013.

WOOLDRIDGE, J. M. **Econometric Analysis of Cross Section and Panel Data**. Cambridge: MIT Press, 2002, 752 p.

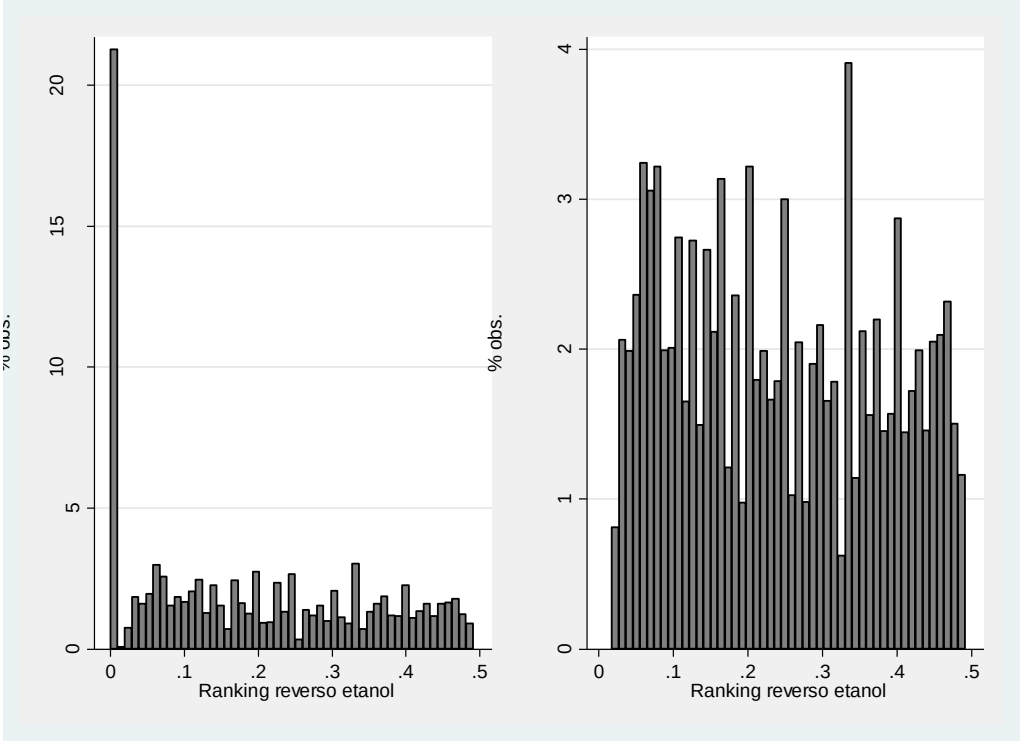
Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP, 2012. **Pesquisa de preços de combustíveis em postos de revenda - Metodologia utilizada para realização da pesquisa de preços no âmbito do Levantamento de Preços e de Margens de Comercialização de Combustíveis da ANP - Vigência: 12/08/2010 a 11/08/2012**. Disponível: www.anp.gov.br/?dw=41567. Acesso em 08 mar. 2013.

APÊNDICE

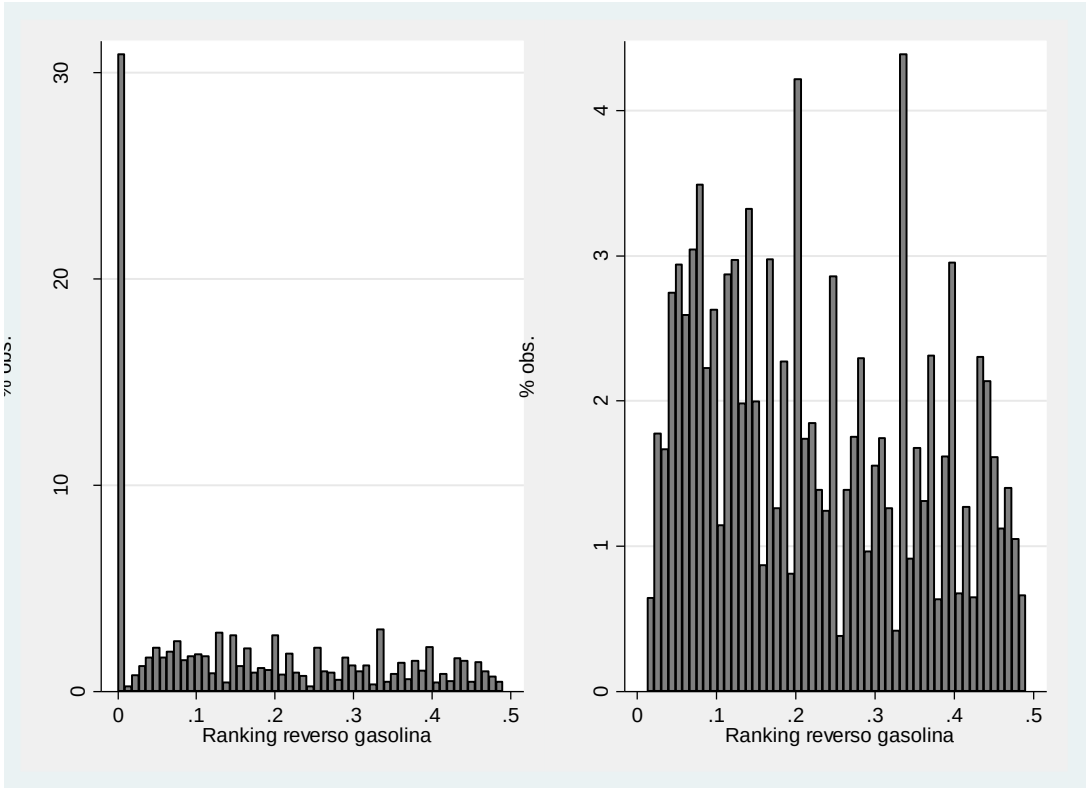
Apêndice 1 – Histogramas de *ranking reverso*.



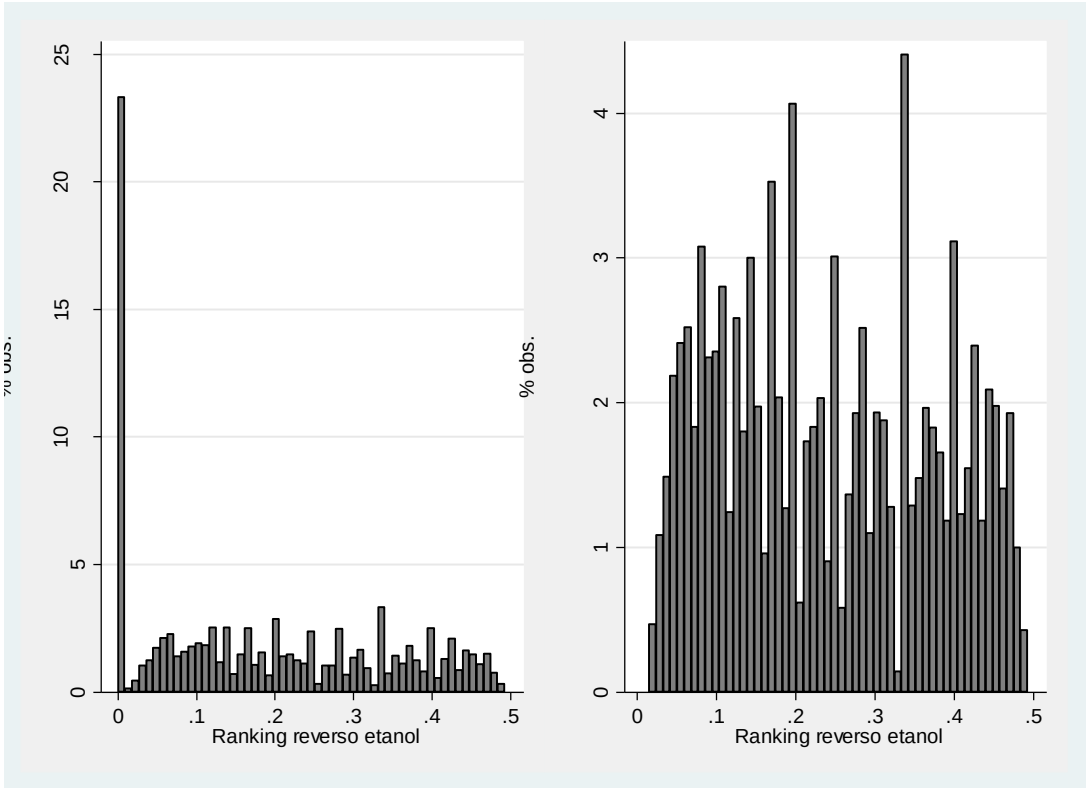
Apêndice 1.1 – *Ranking reverso* gasolina (0,0 a 1,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



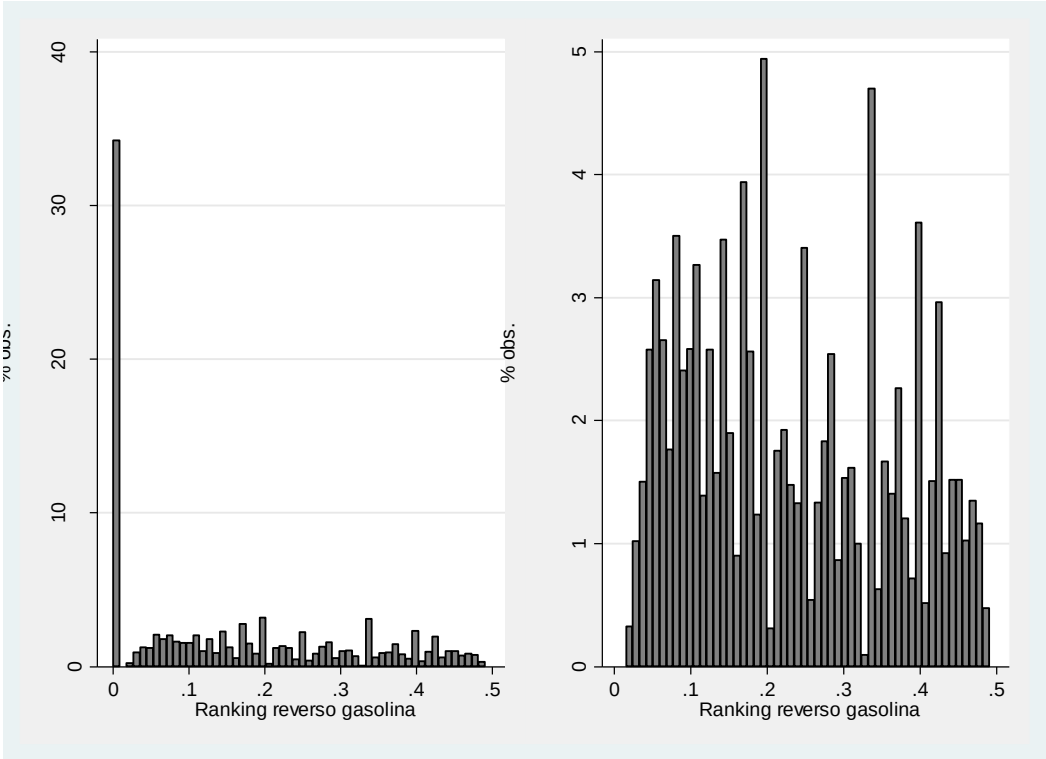
Apêndice 1.2 – *Ranking reverso* etanol (0,0 a 1,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



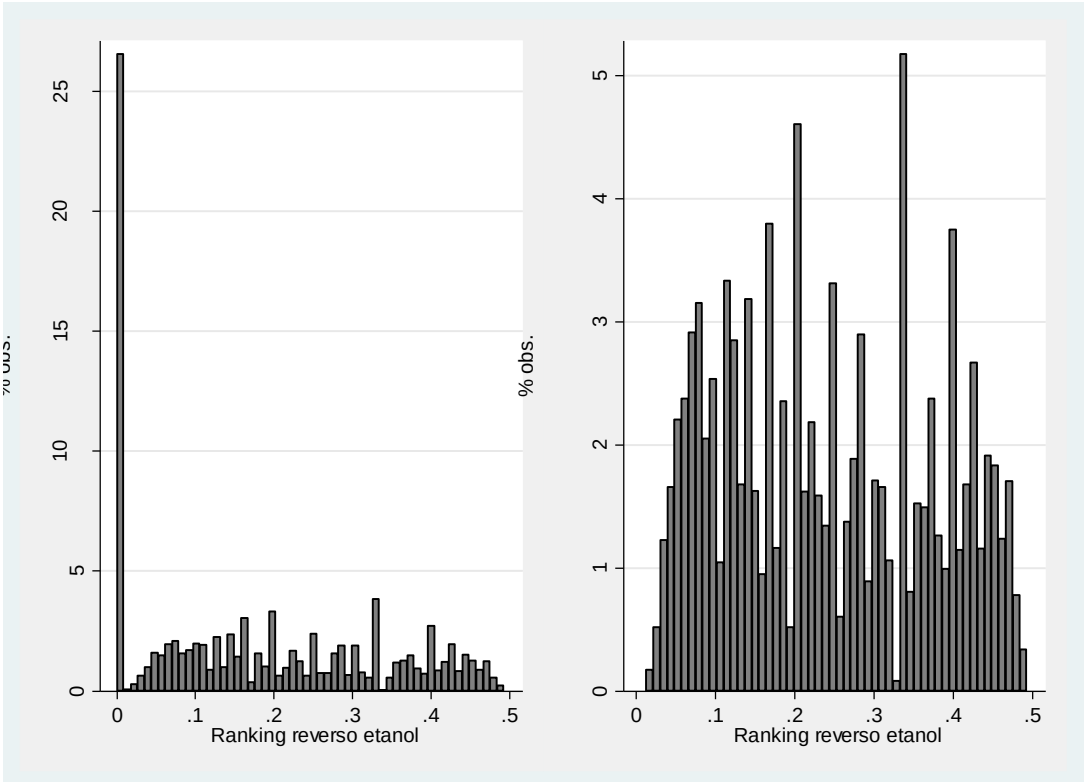
Apêndice 1.3 – *Ranking reverso* gasolina (1,0 a 2,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



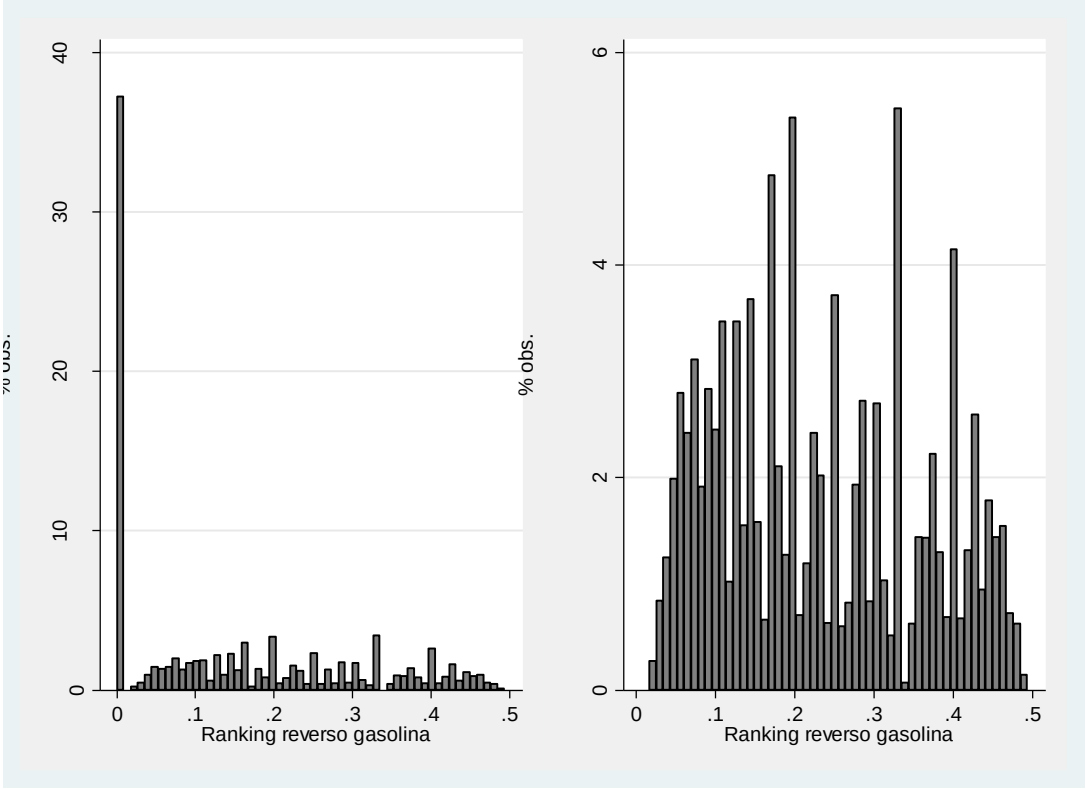
Apêndice 1.4 – *Ranking reverso* etanol (1,0 a 2,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



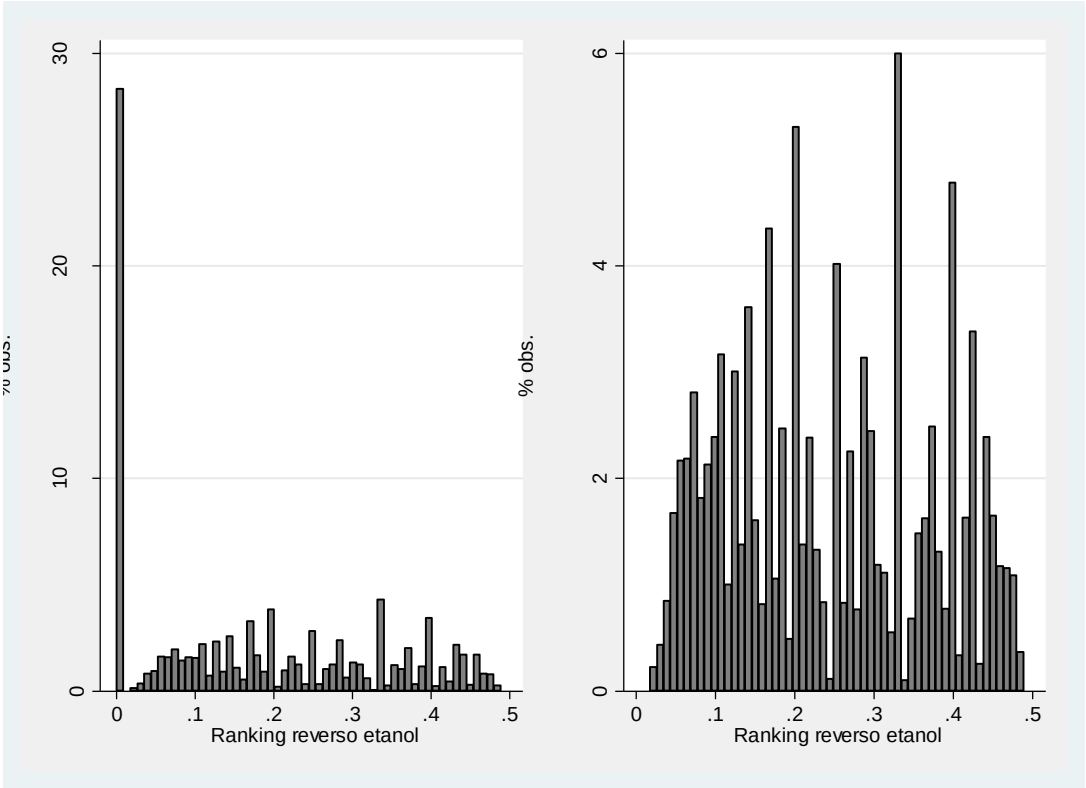
Apêndice 1.5 – *Ranking reverso* gasolina (2,0 a 3,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



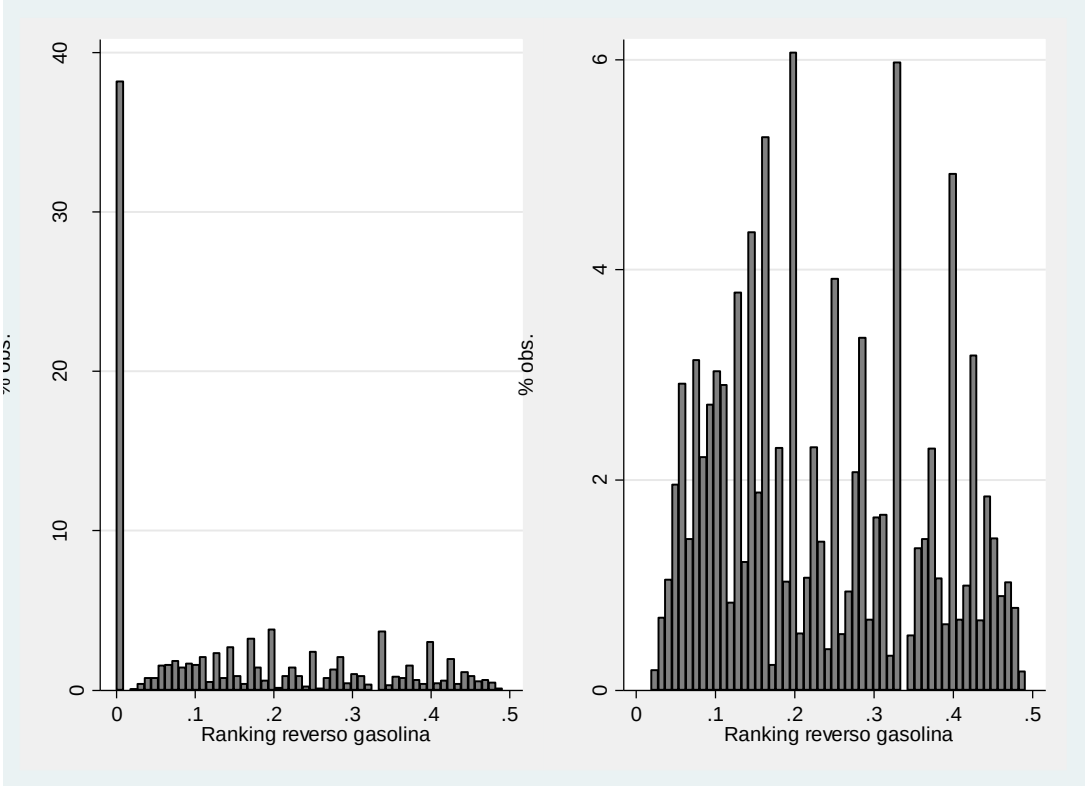
Apêndice 1.6 – *Ranking reverso* etanol (2,0 a 3,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



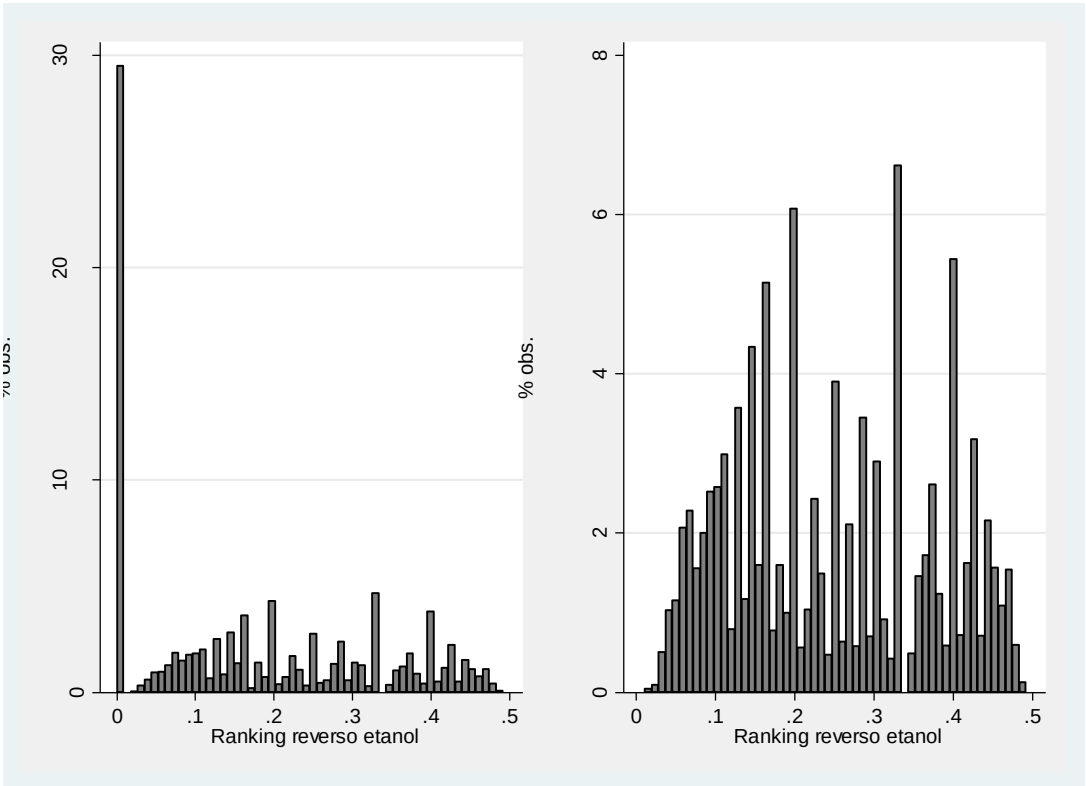
Apêndice 1.7 – *Ranking reverso* gasolina (3,0 a 4,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



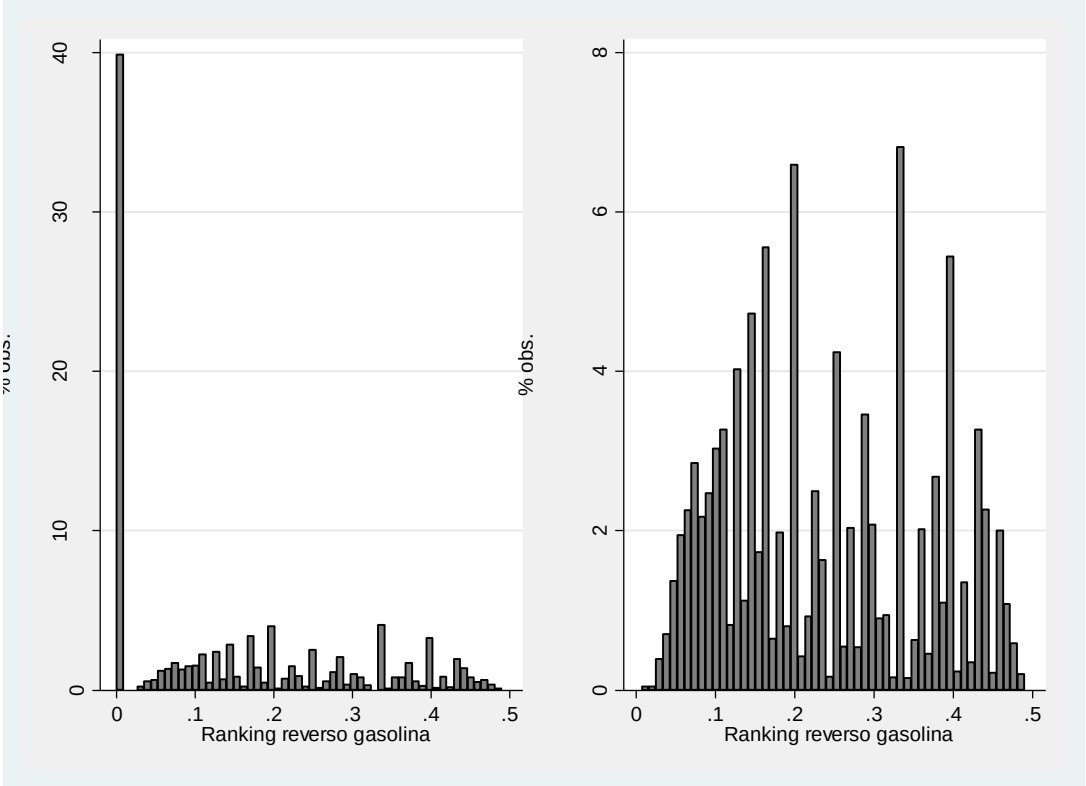
Apêndice 1.8 – *Ranking reverso* etanol (3,0 a 4,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



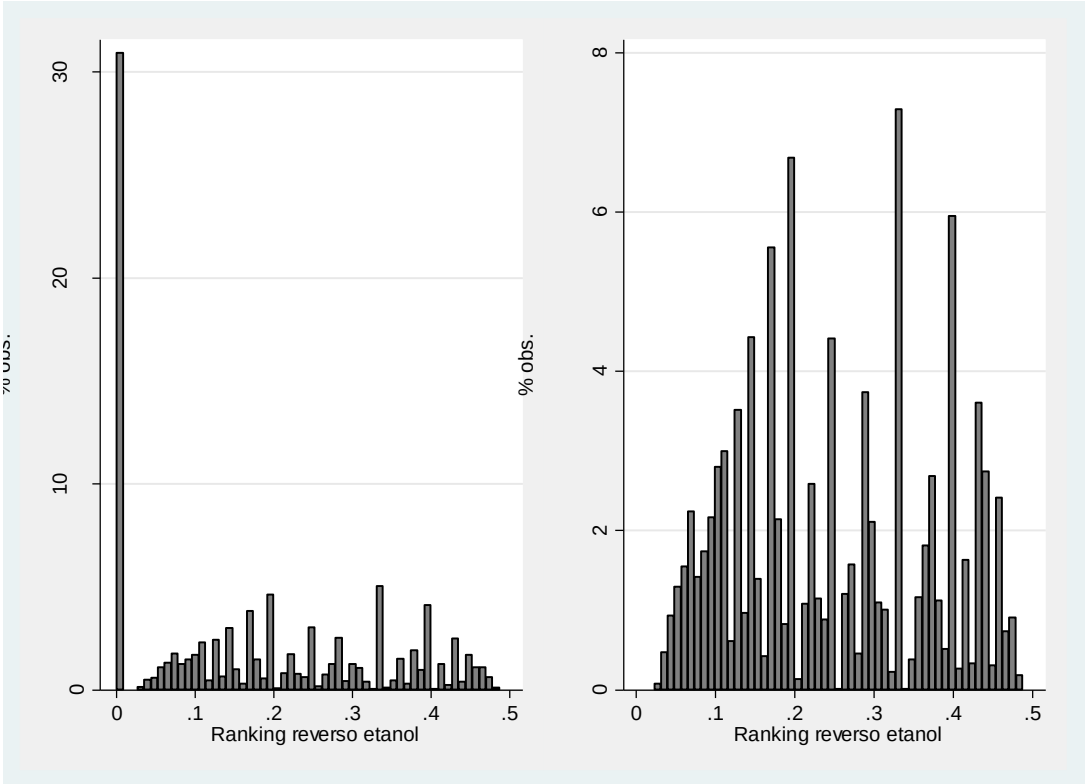
Apêndice 1.9 – *Ranking reverso* gasolina (4,0 a 5,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor



Apêndice 1.10 – *Ranking reverso* etanol (4,0 a 5,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor.



Apêndice 1.11 – *Ranking reverso* gasolina (5,0 a 6,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor.



Apêndice 1.12 – *Ranking reverso* etanol (5,0 a 6,0 km). Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. Elaboração: Autor.

ANEXOS

Anexo 1 – Número de observações por bandeira (*dummies*).

Bandeira	Nº de observações	%
AGIP	796	0,56
ALE	2.208	1,56
ASTER	2.593	1,83
BRANCA	50.541	35,69
BREMEN	88	0,06
CENTROSUL	273	0,19
COSAN	1.814	1,28
ESSO	13.418	9,48
HUDSON	273	0,19
IPIRANGA	17.486	12,35
LIQUIGAS	28	0,02
MASTER	17	0,01
MEG	305	0,22
PETROBRAS	21.482	15,17
PETROFORTE	271	0,19
PETROSUL	657	0,46
REPSOL	770	0,54
RUFF	60	0,04
SALEMCO	172	0,12
SHELL	24.852	17,55
TEXACO	3.021	2,13
TM	472	0,33