

Custos de Procura e Dispersão de Preços de Gasolina e Etanol no Mercado Brasileiro de Combustíveis: Evidências do Município de São Paulo

Search Costs and Price Dispersion of Gasoline and Ethanol in the Brazilian Fuel Market: Evidence from São Paulo City

Camila de Moura Vogt*

Cláudio Ribeiro de Lucinda**

Abstract: A existência de custos de procura é um fenômeno muito comum e pouco estudado no Brasil. O presente estudo busca avaliar a importância dos custos de procura na dispersão de preços de gasolina e de 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 na segunda é descrita a dispersão na área. Os dados dos preços de varejo são coletados semanalmente e a série histórica foi cedida pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. Os resultados das estimações mostram que, apesar da dispersão, possivelmente a competição local é mais relevante para o entendimento da dispersão de preços do que os custos de procura.

Keywords: Concorrência. Dispersão de preços. Custo de procura.

Resumo: Search costs is a common phenomenon, although rarely studied in Brazil. 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 and the time series was provided by the *Agência Nacional do Petróleo, Gás Natural e Biocombustíveis*. The results show local competition characteristics might be more relevant to understand price patterns than search costs.

Palavras-chave: Competition. Price dispersion. Search cost.

JEL Classification: C23; D83.

* Mestra em Economia Aplicada pela Universidade de São Paulo (USP). E-mail: vogt.camila@gmail.com

** Doutor em Economia de Empresas pela Fundação Getúlio Vargas (FGV-SP). Professor associado da Faculdade de Economia e Administração de Ribeirão Preto da Universidade de São Paulo (USP). E-mail: claudiolucinda@fearp.usp.br

1 Introdução

Mensurar como a informação irá afetar as estratégias das firmas e consumidores é um campo recente na literatura econômica. 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. Este trabalho pretende estudar o impacto do custo de procura nos preços de varejo de combustíveis no município de São Paulo e mensurar seus efeitos na dispersão.

Essa opção foi feita principalmente porque se aproxima das justificativas utilizadas no estudo de Chandra e Tappata (2011):¹ custos marginais dos estabelecimentos muito próximos; preços dos produtos acessíveis a todos os consumidores; e disponibilidade de dados semanais sobre preços e localização geográfica do varejo de combustíveis brasileiro.

Toda a base de dados deste 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 entre o período de 2001 a 2012.

Utilizando-se os dados dos preços e a localização dos estabelecimentos, é feita uma estimativa a fim de se mensurar qual a área de mercado relevante dos postos de combustíveis para a análise da dispersão. Utilizando-se como base o trabalho de Davis (2005), que analisa o impacto da concorrência a partir das informações correspondentes a estruturas de mercado, se verifica 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. Também é criado 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 cria uma amostra de controle, na qual o custo de informação é zero, permitindo a comparação com vizinhanças distantes onde há custos de informação.

1 Chandra e Tappata (2011), ao analisarem a dispersão de preços nos postos de combustíveis dos Estados Unidos, citam as seguintes motivações para a análise desse mercado: a) demanda inelástica; b) impossibilidade de armazenamento; c) custos marginais parecidos; d) ausência de restrição de capacidade do produto; e e) preços facilmente disponíveis aos consumidores.

Assim, após se verificar qual o mercado relevante para o estudo de dispersão no município de São Paulo e se traçar as vizinhanças dos postos de combustível, é calculado o índice de dispersão por proximidades. Empregando-se o modelo teórico e os resultados do índice para a amostra de controle, é, então, possível traçar qual a área ótima de procura para o consumidor, considerando-se a dispersão de preços, e qual a influência da informação no mercado de varejo de combustíveis.

2 Revisão de Literatura

A revisão de literatura está dividida em duas seções. Na primeira são discutidos referenciais teóricos de modelos de dispersão de preços. Já na segunda são apresentadas as metodologias empíricas que serão utilizados para a análise dos dados do trabalho.

2.1 Modelos Teóricos de Dispersão de Preços

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 da informação. Nele é descrito 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. É incluído, no custo do produto, o custo da informação, considerado como o valor necessário para que os consumidores sejam informados dos diversos preços. Dessa forma, dependendo de como se comporta esse custo nos mercados, os preços serão diferentes entre as firmas e isso irá ultrapassar as características do produto.

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 precisa 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 mistas, com livre entrada e saída de empresas. No caso de estratégias puras, se considera que existe um equilíbrio de Nash, ou seja, nenhuma firma irá lucrar mais mudando sua estratégia de preços unilateralmente. Já no segundo caso, se destacam os modelos de Salop e Stiglitz (1977) e Diamond (1971), que trabalham com modelos de estratégias mistas. Nesses modelos se inclui a variável informação, que garante que seja possível que as firmas mudem de estratégia dependendo quanto saibam sobre a ações das firmas competidoras.

Também se observam na literatura as buscas sequencial e não sequencial. Para o modelo não sequencial de dispersão (VARIAN, 1980), o consumidor decide entre comprar da loja com o menor preço, incorrendo no custo 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, 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) ou a compra do preço mínimo observado. Na busca sequencial, considera-se que não se está lidando com um ambiente com um tamanho fixo, mas sim com uma procura de forma sequencial, na qual o número de buscas é uma variável aleatória com distribuição geométrica.

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 ao material 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 . Considera-se ainda que o número de informados I , e de 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 $U = M/n$. Cada loja vai possuir uma função de densidade que indica a probabilidade que tem de cobrar o preço p .²

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 idêntica de custo médio e decrescente. Espera-se que as entradas de firmas ocorram até os lucros econômicos chegarem a zero. Para o autor, a randomização dos preços entre as firmas é de interesse para a discriminação de preços entre os consumidores informados e os desinformados. 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.

2 Varian (1980) utiliza a função de probabilidade $(1 - F(p))^{n-1}$.

Baye *et al.* (2006) possuem 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. Isso 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, conseqüentemente, à 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. 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 apontam Chandra e Tappata (2011).

Chandra e Tappata (2011) propõem um teste sobre os impactos da informação no mercado de postos de combustíveis americano a partir da comparação, ao longo do tempo, dos preços entre os estabelecimentos, utilizando, para isso, grupos de controle para informação imperfeita. O modelo teórico assume uma série de suposições baseadas em modelos clássicos como Varian (1980). Os autores consideram um mercado homogêneo com 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 v , sendo o lado da demanda caracterizado por consumidores com demanda inelástica igual a v .

Os consumidores são representados como uma fração λ com distribuição entre 0 e 1. Sendo que possui custo de procura nulo quando a fração é igual a zero, e custo de procura 1 quando há algum custo. 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.³ A proporção de consumidores que opta pela informação irá definir a intensidade de busca desse mercado.

As condições de equilíbrio do modelo serão: a) para a intensidade⁴ $\mu \in [\lambda, 1]$; b) as estratégias das firma precisam seguir um equilíbrio de Nash em estratégias

3 Conforme Chandra e Tappata (2011), as estimações de modelos de busca sequencial são similares aos de busca não sequencial.

4 A intensidade é definida pela proporção de consumidores informados.

mistas; c) a intensidade de procura do mercado μ^* 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 da equação 1:

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

Considerando-se que o preço que maximiza a função é tal que $p \in [p^* = \frac{cn\mu + (1-\mu)v}{1+(n-1)\mu}, v]$. 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 uma função não monotônica de intensidade de procura associada aos ganhos de procura. Em outras palavras, quanto o consumidor irá ganhar pela diferença entre o intervalo do preço máximo e o preço mínimo cobrado pelas firmas em função da quantidade de informação. A equação é definida pelos *gains from search* (GS):

$$GS = E[p - p_{min} | \mu; c, v, n] = \int_{p^*}^v p [1 - n[1 - F(p; c, v, n)]^{n-1}] dF(p; c, v, n) \quad (2)$$

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

Conforme o modelo, a dispersão por si só não pode ser usada para prever a intensidade de pesquisa ou o nível de competição no mercado. Para isso, deve-se definir como a intensidade de busca maximiza o GS para um determinado custo de produção e número de firmas. A relação entre a intensidade de busca e a dispersão de preços é positiva quando $\mu^* < \hat{\mu}$ e negativa 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.

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 1) tendo a intensidade de procura como dada e compara os ganhos de procura (equação 2) com seus gastos.

Em relação ao aumento de consumidores, não é possível considerar o efeito total de mais consumidores informados e desinformados no preço médio já que mais consumidores por meio da procura 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 à hipótese do consumidor decidir antecipadamente entre comprar da loja com o menor preço com o custo 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 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 considerar 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 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.

2.2 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, ou 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 pequenas evidências econométricas. 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.

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 contabiliza 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, h , sendo m o índice que representa o mercado, t , o tempo e $g=1...G$, os 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 , se constrói a seguinte equação:

$$Cinema_{hmt}(a,b) = \sum_{\Gamma_{hmt}(a,b)} Vizinhas_{gmt} \quad (3)$$

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 $rival_{hmt} = rival_{hmt}(0;0,5); rival_{hmt}(0,5;1,0) \dots rival_{hmt}(25;30)$. Então é definida a equação que irá estimar a variação no preço:

$$p_{hmt} = \beta_0 + \alpha_1 rede_{hmt} + \alpha_2 rival_{hmt} + x_{hmt} + \mu_h \beta + \epsilon_{hmt} \quad (4)$$

A variável x é um vetor de controle e as variáveis $rede$ e $rival$ são o somatório de cinemas 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-se 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 abordar os estudos que efetivamente mensuram a dispersão de preços. 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 qual proporção de tempo que um vendedor mantém a estratégia de cobrar o preço mais alto em períodos consecutivos. Espera-se que as firmas randomizem suas posições com relação aos preços ao longo do tempo, assim chegariam a 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 para estabelecimentos mais distantes.

Para isso, é utilizada uma proporção em que dois postos (i, j) sobre o tempo T_{ij} 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:

$$r_{it} = \frac{1}{T_{ij}} \sum_{t=1}^{T_{ij}} I_{(pjt > pit)} \quad (5)$$

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 a uma maior dispersão. O estudo expõe que os consumidores deveriam procurar menos quando os preços sofrem choque e aumentam rapidamente.

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.

3 Metodologia

A presente seção se divide em três partes: a primeira descreve os dados utilizados, a segunda aborda a metodologia de localização dos estabelecimentos e como são traçadas as distâncias e a terceira traz as estimações para um mercado relevante e a mensuração da dispersão.

3.1 Base de Dados

A base de dados foi elaborada em duas etapas. A primeira consiste no levantamento de dados em painel dos preços praticados pelos postos de combustíveis no município de São Paulo durante os anos de 2001 a 2012. Já na segunda etapa é feito o georreferenciamento das informações, a fim de identificar as coordenadas geográficas de cada um dos estabelecimentos levantados na etapa anterior.

3.1.1 Levantamento de Preços

Os dados dos preços praticados pelos postos foram cedidos pela ANP que realiza uma coleta semanal de preços de varejo desde 2001 (AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, 2013b).

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.⁵

Para o estudo do mercado de postos de combustíveis são analisados os dados de preços da ANP referentes à cidade de São Paulo (AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, 2013b). A escolha da região dos dados se deu em função da prática de cartéis atingirem principalmente municípios menores.

De acordo com a Secretária de Defesa Econômica (SDE) de Minas Gerais (SDE/MG, 2011), denúncias de prática de associação de cartel dos postos respondem por um terço do total das denúncias recebidas pelo órgão. Conforme o Conselho Administrativo de Defesa Econômica (Cade), entretanto não há nenhum registro de processo de cartel em andamento no município de São Paulo

5 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.

(CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA, 2014).⁶ Além disso, a própria escala do município é tal que a manutenção de esquemas colusivos em todo o período analisado é pouco plausível. Alguns estudos, entretanto, indicam que a prática de cartel é uma hipótese empiricamente plausível em determinados períodos. Em um estudo para o estado de São Paulo durante o ano de 2004, é observada a variabilidade de preços, que pode ser explicada por colusão no estudo de Nunes e Gomes (2005). Já Vasconcelos e Vasconcelos (2008) observam que, durante o aumento de preços no ano de 2006, registrou-se um comportamento de possível coordenação que indicaria cartel. No entanto, não se constatou a hipótese de menor variância para os preços praticados, o que reforçaria a existência de coordenação.

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

- a) razão social do posto;
- b) local;
- c) bandeira;
- d) preço de revenda da gasolina;
- e) preço de distribuição da gasolina;
- f) preço de revenda do etanol;
- g) preço de distribuição do etanol.

As variáveis de preço de distribuição da gasolina e preço de distribuição do etanol não são 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 atrelada à distribuidora de combustíveis, e “bandeira branca”, quando o posto não está vinculado a nenhuma distribuidora.

O banco de dados inicial trata de mais de 260 mil observações (ver Tabela 1) divididas entre os anos de 2001 e 2012.⁷ São pesquisadas aproximadamente 3.600 razões sociais diferentes entre os períodos.

6 De acordo com dados do Conselho Administrativo de Defesa Econômica (2014), foram julgados, estão aguardando julgamento e estão em trâmite, casos de cartel nas seguintes cidades brasileiras: Florianópolis (SC), Goiânia (GO), Lages (SC), Belo Horizonte (MG), Brasília (DF), Recife (PE), Campinas (SP), Guaporé (RS), Londrina (PR), Manaus (AM), Santa Maria (RS), Teresina (PI), Bauru (SP), Vitória (ES), João Pessoa (PB) e Caxias do Sul (RS).

7 De julho de 2001 a julho de 2012.

Tabela 1 - Descrição dos dados da ANP

Variável	Nº de observações	Média	Desvio padrão	Preço mínimo	Preço máximo
Preço gasolina	260.039	2,224	0,311	1,289	3,399
Preço etanol	257.789	1,274	0,325	0,529	2,899
Preço Distribuição Gasolina	155.451	1,910	0,264	0,825	2,778
Preço Distribuição Etanol	144.815	1,028	0,307	0,300	2,385

Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

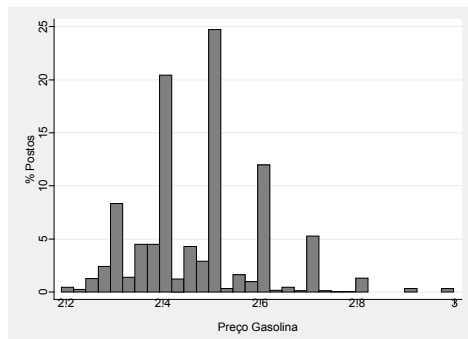
As observações fornecidas pela ANP, em função dos registros serem feitos manualmente, tiveram correções feitas nos endereços. Essa opção foi tomada no intuito de melhorar a efetividade do georreferenciamento e aproximar as informações fornecidas pelo censo demográfico de 2010 do IBGE aos dados dos postos.

O maior preço médio atingido pela gasolina e pelo etanol ocorreu em 2011, ano em que, segundo o Boletim Anual de Preços (AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, 2012), 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 bicompostíveis. 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 é possível que um posto não tenha tido o seu preço coletado nos 133 períodos. Em alguns estabelecimentos, por exemplo, só foram coletados em um período.

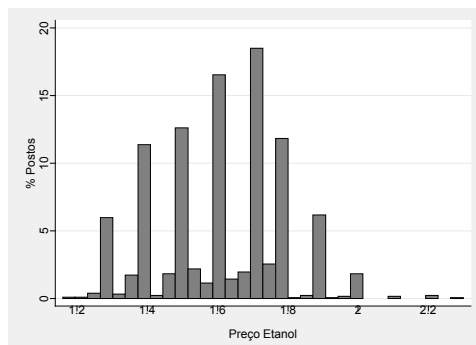
Nos Gráficos 1 e 2 são apresentados os histogramas com a dispersão dos combustíveis gasolina e etanol no período com mais dados. No eixo horizontal constam os preços e no eixo vertical, a porcentagem de postos que praticou determinado preço nesse período. Os histogramas reforçam a ideia de dispersão, uma vez que em nenhum dos momentos foi praticado o mesmo preço por mais de 25% dos postos analisados, seja com relação à gasolina ou etanol.

Gráficos 1 – Histograma de dispersão de preços de venda de gasolina em janeiro de 2004



Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

Gráficos 2 – Histograma de dispersão de preços de venda de etanol em janeiro de 2004



Fonte: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

Apesar de se verificar 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, é necessário saber onde estão localizados os postos. A próxima seção trata do georreferenciamento 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 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 API). Esse sistema, em linguagem, 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 Maps, por meio do programa Google Maps JavaScript API v3, realizou 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 é descrito por Ozimek e Miles (2011), pelo comando *geocode* no Stata. A utilização do comando, entretanto, não foi possível devido a mudanças do 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 aleatória dentro do município de São Paulo.

Com base nos dados de localização foram, então, mapeados os raios de vizinhança entre os estabelecimentos de combustíveis. Para isso, foi utilizado o comando *dismatch* no *software* Stata, que, a partir das informações das coordenadas, traçou as distâncias através da fórmula de Haversine⁸ que cada posto tem do seu concorrente, dentro de um raio limite preestabelecido.

3.2 Modelo Empírico

No modelo empírico, considera-se 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

8 Equação normalmente usada para navegação: a distância de uma esfera é calculada a partir dos pontos das coordenadas. Fórmula utilizada no comando *dismatch*.

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 (2005) para se determinar a extensão geográfica dos mercados relevantes. Já a segunda etapa é baseada no trabalho de Chandra e Tappata (2011). Na Tabela 2 estão divididos os raios de distâncias que serão analisados, utilizando-se como referência os raios de competição de Davis (2005).⁹

Tabela 2 – Raios de competição

Raio	Distância raio
Raio 01	0,0 a 1,0 km
Raio 02	1,0 a 2,0 km
Raio 03	2,0 a 3,0 km
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

Fonte: Elaboração própria.

Dentro do raio 01 estarão os estabelecimentos vizinhos com até 1,0 km de distância entre si, no raio 02, estabelecimentos de 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, busca-se verificar se existe um ambiente competitivo importante nos raios analisados.

O trabalho de Davis (2005) é a base da metodologia para se estabelecer os efeitos do número de estabelecimentos sobre os preços, mas, nesse caso, diferentemente de seu modelo, são acrescentadas *dummies* que identificam a marca do estabelecimento, o tempo e o bairro e os postos não são divididos entre rivais e da mesma rede. São também incluídos, como variáveis de controle, dados do censo 2010 relativos à renda e tamanho da população dos bairros (IBGE, 2010). 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.

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

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 ¹⁰. Desse modo, para todos os postos há uma vizinhança de forma que $Vizinhos_{irt} = Vizinhos_{i1t}$ (0;1,0); $Vizinhos_{i2t}$ (1,0;2,0)... $Vizinhos_{i10t}$ (9,0;10,0). Montados os raios de vizinhança, é estimada a seguinte regressão baseada em Davis (2005):

$$p_{it} = \beta_0 + \beta_r Vizinhos_{irt} + \beta_2 X + \varphi_i + \varepsilon_{it} \quad (6)$$

em que p_{it} é as variáveis de preços, β_0 é o intercepto, $Vizinhos_{irt}$ é a variável relativa ao somatório do número de vizinhos por raio; X é o vetor das variáveis de controle (renda média das famílias por bairro, densidade demográfica por bairro, período e *dummies* de bandeira¹¹ e bairro), φ_i representa um efeito fixo em relação ao tempo e ε_{it} é 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 da equação 5, no qual os preços entre dois postos (i, j) localizados no mesmo raio, sobre o tempo T_{ij} é tal que usualmente¹² $pit \geq pjt$ definindo o *rank* reverso entre os postos i e j como a proporção observada em que $pjt > pit$, conforme a equação 5. O índice revela, dentre os postos que normalmente cobram o maior preço, com que frequência irão cobrar o menor preço.

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 a equação 6; e nos estabelecimentos de “esquina” ou amostra de controle, que foi considerado com uma distância mínima de 100 m.

A amostra de controle permite avaliar uma condição em que 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 6 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.

10 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.

11 As *dummies* de bandeira foram divididas entre as seguintes redes de postos: Shell, BR, Ipiranga, Esso, Texaco, sem bandeira ou rede e outras.

12 O tempo definido em que $pit \geq pjt$ deverá ser no mínimo 50% do período observado. O *rank* reverso, portanto, não poderá superar 0,5.

4 Resultados e Discussão

Nesta seção 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 6, utilizando um modelo em painel de mínimos quadrados ordinários, efeitos fixos¹³ e variáveis instrumentais.

As Tabelas 3 e 4 apresentam os valores dos coeficientes β_r das variáveis $Vizinhos_{it^r}$, sendo $r=(1...10)$. Para as estimações na Tabela 3, 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. Já os resultados da Tabela 4 utilizam variáveis instrumentais, flexibilizando a condição utilizada anteriormente.

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

Tabela 3 – Estimações por raios de proximidade:
mínimos quadrados ordinários e efeitos fixos

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

Fonte: Elaboração própria a partir de dados da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

Nota: ***Significativo a 1%; ** Significativo a 5%; *Significativo a 10%. Erro padrão entre parênteses.

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 3 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 o método de mínimos quadrados e efeitos fixos, respectivamente. Com relação ao etanol, as estimações por mínimos quadrados, com exceção do raio 04, são 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. As variáveis de controle relativas a número de domicílios e renda não apresentaram resultados significativos.¹⁴

Da mesma forma que Davis (2005) opta pela utilização de variáveis instrumentais no intuito de minimizar o problema da endogeneidade, levando em consideração características concorrenciais, como, por exemplo, o fato de que regiões mais prósperas tendem a ter mais estabelecimentos e isso pode afetar o preço, se considera que haja correlação diferente de zero entre o termo de erro e os regressores analisados.

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 o erro e a variável instrumentalizada. O uso de estruturas de mercado com defasagem de quatro 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 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 tentam demonstrar o mercado onde os estabelecimentos estão inseridos.

As estimações nas colunas 3.1 e 3.2 da Tabela 4 mostram o uso de variáveis instrumentais utilizando a quarta defasagem das estruturas de mercado. Cada linha da tabela representa um coeficiente β_i da equação 6.

14 Variáveis de controle renda e densidade demográfica utilizadas por Davis (2005).

Tabela 4 – Estimações de raios de proximidade com variáveis instrumentais

β_r	Variáveis instrumentais (3.1)	Variáveis instrumentais (3.2)
Σ Vizinhos	gasolina	etanol
Σ Vizinhos (raio 01)	-0,169791 (-1,140)	-0,205862 (-1,287)
Σ Vizinhos (raio 02)	0,044549 (0,636)	0,042391 (0,564)
Σ Vizinhos (raio 03)	0,024400 (0,298)	0,057069 (0,631)
Σ Vizinhos (raio 04)	0,274669* (1,741)	0,296874* (1,734)
Σ Vizinhos (raio 05)	-0,444784** (-2,163)	-0,530444** (-2,188)
Σ Vizinhos (raio 06)	0,801099*** (2,638)	0,954054** (2,556)
Σ Vizinhos (raio 07)	-0,811543** (-2,513)	-0,808950** (-2,283)
Σ Vizinhos (raio 08)	0,096471 (0,288)	-0,153851 (-0,390)
Σ Vizinhos (raio 09)	-0,032443 (-0,110)	0,103392 (0,303)
Σ Vizinhos (raio 10)	0,224197 (1,064)	0,222918 (1,009)
R2	-0,001	-0,001
Número de observações	32.751	32.751

Fonte: Elaboração própria a partir de dados da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

Nota: ***Significativo a 1%; ** Significativo a 5%; *Significativo a 10%. Erro padrão entre parênteses.

Pode-se 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-Hansen¹⁵ e o teste de significância conjunta das variáveis, entretanto,

15 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. A rejeição põe em dúvida a validade dos instrumentos.

rejeitam a hipótese nula, mostrando que o conjunto de coeficientes não é significativo. Não é possível, portanto, restringir a área geográfica máxima em que mais de um estabelecimento poderia ter impacto marginal na busca do consumidor e na dispersão de preços. No intuito de restringir o mercado para a análise da dispersão, se optou por utilizar a área até o raio 07.

4.2 Dispersão Temporal de Preços

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.

Considera-se 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 e 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,¹⁶ calcula-se então o *rank* reverso conforme a equação 5, ou seja, a proporção no tempo em que $pjt \geq pit$.

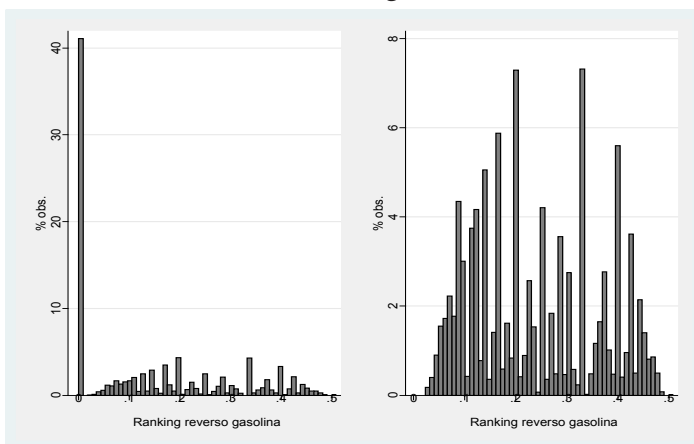
Dessa forma, o *rank* reverso será a proporção de tempo em que o posto que normalmente pratica o preço maior irá ter o menor preço. Em outras palavras, entre os postos que normalmente cobram o maior preço, com que frequência irão mudar de posição cobrando o menor preço.¹⁷

Os Gráficos 3 e 4 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.

16 Em mais de 50% do tempo.

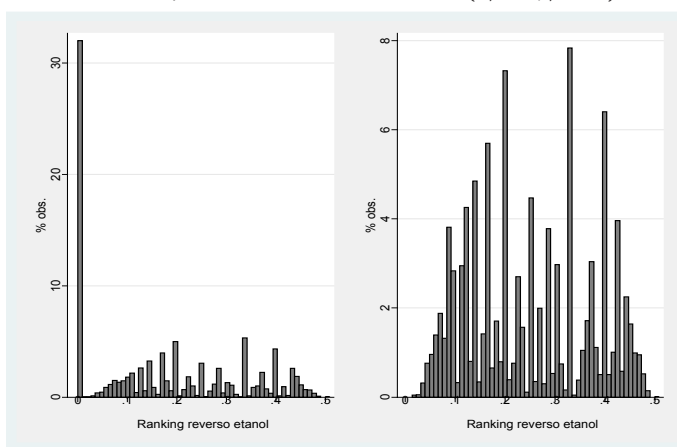
17 As estimações consideraram apenas pares de postos que possuíam uma frequência de cinco ou mais períodos observados.

Gráfico 3 – Rank reverso da gasolina (6,0 a 7,0 km)



Fonte: Elaboração própria a partir de dados da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

Gráfico 4 – Rank reverso do etanol (6,0 a 7,0 km)



Fonte: Elaboração própria a partir de dados da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

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 pode-se observar que 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 diferente das observações do estudo de Chandra e Tappata (2011), no qual é utilizada uma base censitária dos preços.

Na Tabela 4, tem-se 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,¹⁸ classificados como “esquina”.

Tabela 5 – Resultados do *rank* reverso

Raio	Combustível	Número de pares	Rank reverso médio	Diferença média de preço (preço de origem e 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

Fonte: Elaboração própria a partir de dados da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

Os estabelecimentos de “esquina” são utilizados como amostra de controle para a busca do consumidor. Considerando que os preços estão expostos na frente

18 Considera-se distância igual ou inferior a 100 m.

dos estabelecimentos, os consumidores terão acesso visual aos preços e poderão fazer a sua escolha pelo menor preço.

Os resultados da Tabela 5 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, em que 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 menor, ou seja, postos mais próximos têm maior dispersão, mas a diferença entre os preços é menor.

Para uma verificação mais robusta da influência da busca do consumidor, se assume 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, $F_0(r)$. Se busca do consumidor tem influência então $F_0(r) < F_1(r)$, ou seja, a dispersão de preços deve ser menor em estabelecimentos que não tenham assimetria de informação.

Utilizando-se o teste Kolmogorov-Smirnov (K-S), no qual a hipótese nula $H_0: F_0(r) = F_1(r)$, e hipótese alternativa $H_1: F_0(r) \neq F_1(r)$, pode-se 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, usam-se os resultados da equação:

$$r_{it} = \beta_0 + \beta_1 I_{it} + \beta_2 X_{it} + \varepsilon_{it} \quad (7)$$

em que 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 o censo 2010 (IBGE, 2010, 2011). 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.

No entanto, os resultados associados aos valores de *rank* reverso (Tabela 5) nos mostram que os postos de controle, ou os postos “de esquina”, apresentam um coeficiente de *rank* reverso maior do que os dos postos não vizinhos, conforme o Quadro 1:

Quadro 1 – Resultados estimações rank reverso: postos de “esquina” versus mercado

Raio < 1,0 km			
			<i>p-value</i>
Gasolina	Dummy Iij	0,009	0,001
	0	0,063	0,000
	1	-0,018	0,094
	(K-S)	0,063	0,000
Etanol	Dummy Iij	0,028	0,000
	0	0,099	0,999
	1	0,000	0,000
	(K-S)	0,099	0,000

Fonte: Elaboração própria a partir de dados da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2013b).

Nota: $H_{\alpha}:F_0(r_{ik})=F_1(r_{ij})$, em que $d_{ik}<0,1$ km e $0,1$ km $<d_{ij}<1,0$ km; $H_{\alpha}(0):F_0(r_{ik})>F_1(r_{ij})$; $H_{\alpha}(1):F_0(r_{ik})<F_1(r_{ij})$; $H_{\alpha}(KS):F_0(r_{ik})\cong F_1(r_{ij})$.

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 diferenciação 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 na amostra estudada, ou seja, é possível que a diferenciação de produtos, discriminação de preços e choques de oferta tenham um papel mais relevante. Esse resultado pode demonstrar que aspectos concorrenciais podem ter uma influência significativa sobre o comportamento de preços.

Na literatura nacional, o mercado de varejo de combustíveis no Brasil é um campo com ampla pesquisa sobre aspectos concorrências e de comportamento de preços. Questões relativas à dinâmica concorrencial entre distribuidores, varejistas e consumidores são exploradas através da perspectiva de transmissão de preços. Uchôa (2008) defende, através de testes de cointegração entre séries de tempo, que existe o ajustamento assimétrico de preços entre atacados e distribuidores. A mesma metodologia em séries desagregadas (SILVA *et al.*, 2014) para o Brasil, entretanto, mostra ajustamento simétrico, o que levanta dúvidas sobre o impacto dos aspectos concorrenciais entre atacado e varejista para os consumidores finais.

Outros aspectos concorrenciais também são abordados na literatura, com ênfase para aspectos de coordenação de preços entre varejistas, como nos estudos de Nunes e Gomes (2005), Sampaio e Sampaio (2013) e Silva *et al.* (2014).

5 Considerações Finais

Com base na metodologia de Davis (2005), não foi possível estabelecer o limite geográfico máximo do impacto marginal de um concorrente sobre os preços de varejo de gasolina e etanol. Os resultados de *rank* reverso, entretanto, 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 queira pesquisar qual o melhor preço não precisará 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 (COMPANHIA DE ENGENHARIA DE TRÁFEGO, 2013), 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 de dispersão de preços considerando a procura do consumidor. 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. Desse modo, os consumidores, ao pesquisarem os preços, devem levar em consideração a chance de postos de esquina variarem constantemente sua posição de menor ou maior preço frente aos concorrentes próximos.

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.

Em resumo, este estudo conclui que 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 considere a distribuição de consumidores informados, não há dispersão significativa relacionada à busca do consumidor. Se considerados os estudos sobre o mercado concorrencial de varejo, esse resultado pode reforçar a existência de coordenação de preços entre postos vizinhos.

Referências

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS. *Boletim Anual de Preços 2012: preços de petróleo, gás natural e combustíveis nos mercados nacional e internacional*. Rio de Janeiro: 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. 2013a.

_____. SLP- Sistema de Levantamento de dados – 2001-2012. Disponível: <<http://www.anp.gov.br/preco>>. Acesso em: 08 mar. 2013b.

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.

CHANDRA, A.; TAPPATA, M. Consumer search and dynamic price dispersion: an application to gasoline market. *Rand Journal of Economics*, Santa Monica, USA, v. 42, n. 4, p. 681-704, 2011.

COMPANHIA DE ENGENHARIA DE TRÁFEGO. *Pesquisa de monitoração da fluidez: desempenho do sistema viário principal: volume e velocidade – 2012*. São Paulo, mar. 2013. Disponível em: <www.cetsp.com.br/media/228058/2012%20%20volumes%20e%20velocidades.pdf>. Acesso em: 15 out. 2013.

CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. *Cadernos do Cade: Varejo de Gasolina*. Brasília, DF: CADE, 2014. Disponível em: <www.cade.gov.br/aceso-a-informacao/publicacoes-institucionais/dee-publicacoes-anexos/cadernos-do-cade-varejo-de-gasolina.pdf>. Acesso em: 10 set. 2017.

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- 707, 2005.

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

IBGE. *Base de Informações do Censo Demográfico 2010: resultados do universo por setor censitário*. Rio de Janeiro: Documentação do arquivo, 2011.

_____. *Censo Demográfico 2010*. Brasília, DF: IBGE, 2010. Disponível em: <www.censo2010.ibge.gov.br>. Acesso em: 15 out. 2013.

- 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.
- NUNES, C.; GOMES, C. Aspectos concorrenciais do varejo de combustíveis no Brasil. In: ENCONTRO NACIONAL DE ECONOMIA, 33., 2005, Natal. *Anais...* Natal: Anpec, 2005.
- 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.
- 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.
- SAMPAIO, R. M. B.; SAMPAIO, L. M. B. Competição entre postos de gasolina em pequenos municípios brasileiros: uma análise a partir de um modelo de entrada empírico. *Economia Aplicada*, Ribeirão Preto, v. 17, p. 579-598, 2013.
- SECRETARIA DE DEFESA ECONÔMICA DE MINAS GERAIS. Prestação de contas ordinária anual relatório de gestão do exercício de 2010. Brasília, 2011. Disponível em: <<http://portal.mj.gov.br>>. Acesso em 08/03/13.
- SHILONY, Y. Mixed pricing in oligopoly. *Journal of Economic Theory*, Ithaca, v. 14, n. 2, p. 373-388, 1977.
- SILVA, A. S. et al. Symmetric transmission of prices in the retail gasoline market in Brazil. *Energy Economics*, v. 43, p. 11-21, 2014.
- STAHL, D. O. Oligopolistic pricing with sequential consumer search. *American Economic Review*, Pittsburg, v. 79, n. 4, p. 700-712, 1989.
- STIGLER, G. The economics of information. *Journal of Political Economy*, Chicago, v. 69, n. 3, p. 213-225, 1961.
- UCHOA, C. F. A. Testando a Assimetria nos preços de gasolina brasileira. *Revista Brasileira de Economia*. Rio de Janeiro, v. 62, n. 1, p. 103-117, 2008.
- VARIAN, H. R. A model of sales. *American Economic Review*, Pittsburg, v. 70, n. 4, p. 651-659, 1980.
- VASCONCELOS, S. P.; VASCONCELOS, C. R. F. Análise do comportamento estratégico em preços no mercado de gasolina brasileiro: modelando volatilidade. *Análise Econômica*, v. 26, p. 207-222, 2008.

Recebido em: 29/09/2014.

Aceito em: 05/03/2017.