



TEXTO PARA DISCUSSÃO Nº 18

**ESTIMAÇÃO DE PREÇOS DE REFERÊNCIA PARA O SETOR PÚBLICO
COM O USO DE PEQUENAS AMOSTRAS**

Gabriel Lyrio de Oliveira*

André Luis Squarize Chagas†

São Paulo, agosto de 2020

* Doutorando em Teoria Econômica pela FEA/USP.

† Professor do Departamento de Economia da FEA/USP.

Resumo: Este artigo apresenta um novo método de estimação para o preço de referência de um produto, para fins de balizar decisões do poder público relacionadas ao início do processo licitatório. São discutidas as vantagens do novo método em relação a estimação tradicionalmente realizada, via média em pequenas amostras observadas em um determinado período. O estimador desenvolvido decompõe o preço de um bem em um efeito médio de seu setor, de período, de seus fornecedores potenciais e do próprio bem. Com o novo procedimento, não se faz mais necessária a pesquisa de preços de todas as empresas conhecidas em um mesmo período, o que resulta também em maior liberdade no planejamento da coleta de preços, com significativa redução de custos no processo de coleta. Para validação do procedimento, realizaram-se testes comparativos entre o estimador proposto (estimador econômico) e o estimador tradicional (estimador estatístico) com dados reais de coleta de preços de obras públicas. Além deste, outro teste foi construído ao gerar um ambiente simulado, em que se conhece o verdadeiro preço de mercado. Ambos os testes demonstraram que o estimador desenvolvido possui menor variância e menor desvio médio em relação ao preço de mercado, todas as vezes em que não se alcançaram todas as empresas existentes.

ÍNDICE

1.	INTRODUÇÃO	1
2.	PREÇOS DE MERCADO	5
2.1	PREÇOS DE MERCADO: ABORDAGEM ESTATÍSTICA	5
2.2	PREÇOS DE MERCADO: ABORDAGEM ECONÔMICA	6
2.2.1	Empresa de Insumo Único	7
2.2.2	Empresa com Múltiplos Insumos de Produção	8
2.2.3	Modelo Empírico para o Estimador Econômico	10
3.	DADOS E METODOLOGIA	12
3.1	DADOS	12
3.2	MÉTODO	13
4.	RESULTADOS	15
4.1	EXPERIMENTO COM PREÇOS SIMULADOS	15
4.2	DISTÂNCIA ENTRE ESTIMADORES ECONÔMICO E ESTATÍSTICO	20
4.3	EXPERIMENTO COM PREÇOS REAIS	22
5.	CONCLUSÃO	26
6.	REFERÊNCIAS	28
	ANEXO 1: MODELO ECONOMÉTRICO	30

ÍNDICE DE TABELAS

Tabela 1: Descrição dos 147 Setores _____	12
Tabela 2: Estatísticas Descritivas das 100 Simulações _____	17
Tabela 3: Desvios Absolutos do Estimador Econômico e Estatístico em relação ao Preço de Mercado Simulado _____	18
Tabela 4: Distância entre Estimador Econômico e Estatístico _____	21
Tabela 5: Desvio Absoluto em relação ao Preço de Mercado: Estimador Econômico vs. Estatístico ____	23
Tabela 6: Resumo de Resultados _____	24

ÍNDICE DE FIGURAS

Figura 1: Distribuição do Erro – Estimador Econômico vs. Estimador Estatístico _____	19
Figura 2: Preços de Referência - Econômico, Estatístico e o Preço de Mercado _____	19

1. INTRODUÇÃO

Nas licitações de obras públicas no Brasil, a prática consolidada é realizar leilões reversos, onde os lances são dados até um limite pré-estabelecido. Para fixar esse valor, o órgão contratante realiza pesquisa de preço de mercado. De fato, todas as compras governamentais necessitam ser precedidas por uma pesquisa de preços de mercado, que balizará a licitação.

A Lei 8666/1993 regulamenta as normas para licitações e compras públicas. Essa lei estabelece que as compras devem ser processadas por meio de um sistema de registro de preços (art.14, item II), precedido por uma ampla pesquisa de mercado (art.14 item V, par. 1). Nem sempre, no entanto, o órgão contratante tem a necessária experiência no levantamento de preços de mercado. Embora, para o senso comum, pareça algo simples, realizar um levantamento de preços exige detalhamentos, cobertura, periodicidade, capacidade de tratamento de dados, entre outros fatores, que escapam ao conhecimento do órgão contratante. A prática consagrou a realização de “três orçamentos” como a forma de suprir informações ausentes no momento de montar o orçamento da licitação. Mas, três observações em um universo desconhecido não conferem segurança alguma sobre a estatística derivada dessa amostra.

A bem da verdade, esse processo de levantamento de preços não apresenta embasamento em nenhuma teoria estatística ou econômica e segue apenas recomendações processuais ou jurídicas não especializadas. Em um momento em que o país como um todo parece redescobrir a necessidade de que decisões públicas sejam embasadas metodologicamente (cientificamente, na linguagem do senso comum) parece oportuno discutir também o processo de obtenção de preços de referência para licitações.

A disponibilidade de preços de referência gera informações fundamentais para o planejamento do comprador, seja ele grande como no caso do Setor Público e de grandes empresas, seja para o consumidor final. Para o caso do Setor Público, os preços de referência auxiliam na elaboração de licitações servindo como parâmetros do que são preços aceitáveis a serem pagos, evitando assim o desperdício de recursos públicos (SANTOS, 2015). Para contratos públicos relevantes, caso das obras públicas, existe um

acompanhamento mais sistemático por parte de órgãos interessados, que leva a contratação de tabelas de preços específicas. A legislação inclusive já disciplina a utilização dessas tabelas, como é o caso do Sistema Nacional de Pesquisa de Custos e Índices da Construção Civil (Sinapi), mantido pela Caixa Econômica Federal (CEF). Obras de infraestrutura e transporte utilizam o Sistema de Custos referenciais de Obras (Sicro), mantido pelo Departamento Nacional de Infraestrutura e Transportes (Dnit). No Estado de São Paulo, existe há mais de 40 anos o Índice de Preços de Obras Públicas, elaborado pela Fundação Instituto de Pesquisas Econômicas (Fipe). Desse índice foi derivada uma série de tabelas de preços de referência de interesse de órgãos específicos.

Tradicionalmente, os preços dessas tabelas são obtidos tomando a média aritmética dos preços pesquisados em entrevistas junto a um grupo de fornecedores, em um determinado mês. Algumas vezes podem ser aplicados critérios simples para tratar preços discrepantes (*outliers*), desconsiderando-os da amostra. Este estimador é denominado como Estimador Estatístico do Preço de Mercado, pois sua base é a teoria estatística.

Na prática é muito difícil, se não impossível, que se conheça e se colem os preços de todos os ofertantes de um determinado bem no mercado. Para o caso hipotético em que se conhecem todos os fornecedores, seria ainda irreal supor que a coleta do preço de todos fosse possível e financeiramente viável de ser realizada em todos os períodos. Este último fato por si só gera a necessidade da estimação de preços, caso se queira um preço de referência sempre atualizado. Instituições que divulgam preços de referência lidam com essa necessidade adotando métodos de estimar os preços que não foram observados, como pode ser visto, por exemplo, na metodologia exposta em Sinapi (2017).

A utilização de dados de pesquisa de preços para preços de referência difere do preço efetivo de compra por vários fatores. Um deles é o denominado “efeito cotação” (SILVA FILHO, LIMA e MACIEL, 2010). O procedimento de coleta é incapaz de simular o processo concorrencial de contratação, resultando em preços médios maiores do que os que serão vistos no processo de contratação. Pelo efeito cotação, tem-se que os preços de referência devem ser considerados como superiores ao esperado para a compra pelo setor público e que a concorrência na fase de licitação é necessária para que o preço de contratação reflita de forma mais precisa o valor de mercado.

Outro problema mais grave associado à coleta de preços decorre da possível variação de preços devida tão somente pela variação de amostra. No caso em que não se conhecem todos os fornecedores de um bem, e para os que são conhecidos pode não se conseguir informação em um determinado mês, o Estimador Estatístico registrará uma variação do Preço de Referência simplesmente pelo fato de a amostra considerada no período mais atual diferir da amostra de fornecedores considerada no último período de pesquisa. Este problema, denominado flutuação de amostra, é tão grave quanto menor for a amostra de fornecedores pesquisados e conhecidos para o bem em questão.

O estimador para os Preços de Mercado apresentado neste trabalho é denominado Estimador Econômico, pois se baseia na teoria econômica de determinação de preços e utiliza toda a informação disponível para o cálculo do Preço de Referência. Por meio de algumas observações dos preços dos fornecedores de cada bem, mesmo que distribuídas ao longo do tempo, e com a classificação por similaridade setorial destes bens, consegue-se estimar qual teria sido o preço de cada bem informado por um dado fornecedor daquele setor em determinado período, mesmo que tal preço não tenha sido pesquisado no período desejado. Ou seja, obtém-se Preços de Referência atualizados para todos os bens, independentemente de quais foram pesquisados no mês desejado.

As condições necessárias para que este estimador seja calculado é que cada produto esteja classificado em um grupo de produtos similares, que compartilhem componentes de custo e de demanda comuns, de forma que observando alguns produtos do mesmo grupo no período desejado, é possível extrair o efeito médio do período. O efeito médio do produto pode ser obtido ao se observarem alguns de seus preços em períodos passados. Da mesma forma, para obter o efeito médio das empresas que se quer considerar como potenciais fornecedores daquele bem, deve-se ter alguns preços destas observados para alguns produtos do mesmo grupo.

Ao serem considerados os efeitos médios dos produtos, das empresas, do setor e dos períodos para cada preço, se incorpora na formação do Preço de Referência um imenso

conteúdo informacional, através de uma estrutura fundamentada em lógica econômica¹. Esse fato se encontra em linha com avanços recentes da área de estimação de Índices de Preços da literatura internacional (CAVALLO e RIGOBON, 2016), e com os avanços computacionais, que permitem cada vez mais tratar um volume maior de dados em menor tempo, incorporando inteligência ao processo.

Diferente das tabelas de preços, os índices nada mais são do que agregações dos preços de referência estimados para representar preços de mercado, em que atribuindo pesos às variações de preços de um determinado grupo de bens, calculam-se indicadores que servem para agentes do setor público ou privados, por exemplo, reajustarem contratos. Constituem, portanto, potenciais aplicações do método aqui apresentado.

O artigo se encontra dividido em uma seção com motivação teórica para o emprego do modelo econométrico apresentado. Em seguida são apresentados os dados utilizados para alguns dos testes comparativos desenvolvidos. Então, apresentam-se a metodologia de estimação e as características do método proposto. Por fim, são apresentados os procedimentos e resultados obtidos em cada teste comparativo entre o desempenho do Estimador Econômico e do Estimador Estatístico, em termos de variância e de desvio em relação ao supostos Preços de Mercado.

¹ Produtos e fornecedores de um mesmo setor, que estão sujeitos a choques similares de custos e de demanda.

2. PREÇOS DE MERCADO

Preços de mercado podem ser encarados como variáveis aleatórias, cuja esperança matemática pode ser apurada a partir de sucessivas observações. Esse é o princípio que embasa uma abordagem estatística para o problema de observação de preços. Por outro lado, o preço de mercado, do ponto de vista econômico, representa o equilíbrio econômico entre oferta e demanda de um determinado bem. Ele equilibra os desejos de compra e venda dos agentes de mercado. Do ponto de vista da empresa, os preços precisam ser suficientes para cobrir seus custos de produção, incluindo o custo de oportunidade do capital.

Na sequência discutimos com um pouco mais de detalhe essas duas abordagens para estimar o preço de mercado.

2.1 PREÇOS DE MERCADO: ABORDAGEM ESTATÍSTICA

Seja p_i o preço de um determinado produto praticado pela i -ésima firma, em um mercado com M firmas. Então, define-se o preço médio de mercado por:

$$\bar{p} = \frac{1}{M} \sum_{i=1}^M \phi_i p_i \quad (1)$$

Onde ϕ_i é o *market share* de cada empresa. Essa fórmula diz que o preço médio de mercado é composto pelo preço de cada empresa ponderado pela frequência com que os consumidores se deparam com os mesmos. Ou seja, se os consumidores consomem mais da empresa A do que da empresa B, por qualquer motivo, o preço da primeira empresa deve ter maior peso no cômputo do preço de mercado.

Na prática, os institutos de pesquisa não observam o *market share* de cada empresa por produto. Essa é uma variável não observável e tentar obtê-la significaria um ganho informacional desproporcionalmente menor em relação ao custo de fazer esse levantamento. O que se faz, então, é atribuir pesos iguais às empresas, de modo que o preço de mercado pode ser definido como:

$$\bar{p} = \frac{1}{M} \sum_{i=1}^M p_i \quad (2)$$

Ou seja, o preço de mercado é a média aritmética dos preços das empresas que o fornecem.

Outra limitação operacional refere-se ao fato de que é impraticável coletar os preços de todas as empresas para operacionalizar o cálculo em (1). Mais uma vez, o custo de tal esforço não compensaria o ganho de informação adicional. O que se faz, então, é obter uma estimativa dessa média com o emprego de técnicas estatísticas.

Um resultado estatístico importante (Lei dos Grandes Números) é que, para uma amostra aleatória suficientemente grande, a média amostral tende à média populacional. Desse modo, o preço médio de mercado é estimado usando o seguinte estimador:

$$\hat{p} = \frac{1}{N} \sum_{i=1}^N p_i \quad (3)$$

Onde N é o número de fornecedores encontrados e entrevistados, sempre menor ou igual a M . Sobre esse estimador, tem-se que:

1. O cálculo do preço médio usando uma amostra é também uma estimativa;
2. A qualidade desse estimador depende da distribuição da população, do tamanho da amostra, da forma de extração dos dados amostrais, entre outros fatores;
3. Denomina-se esse estimador de “estatístico”, pois seu cálculo se baseia na teoria estatística.

2.2 PREÇOS DE MERCADO: ABORDAGEM ECONÔMICA

Como já foi dito, o preço de mercado reflete as condições de oferta e demanda. Ele é a realização de uma dupla coincidência de desejos, representando o preço que o conjunto de consumidores está disposto a pagar pela quantidade vendida do bem, e o preço pelo qual os ofertantes estão dispostos a vender a quantidade de mercado do bem.

Toda a abordagem microeconômica moderna, herdeira da teoria do valor clássica, se baseia em entender como os preços são determinados em diferentes condições de mercado: concorrência perfeita, monopólios, oligopólios, leilões, contratos de curto ou longo prazo, entre muitas outras determinações.

Em linhas gerais, os preços de mercado dependem de vários fatores:

- Pelo lado da demanda:
 - Estrutura de preferências, preços de bens relacionados no consumo, renda, expectativas etc.
- Pelo lado da oferta:
 - Preços de insumos, preços de fatores de produção, tecnologia, expectativas etc.

Do ponto de vista dos ofertantes, o preço de mercado se relaciona com o custo de produção da seguinte forma:

$$p_i = \mu_i c_i \quad (4)$$

Em que c_i é o custo marginal, ou seja, o custo de produzir uma unidade adicional, $\mu_i \geq 1$ é o mark-up, i.e., a margem sobre o custo marginal.

2.2.1 Empresa de Insumo Único

O custo total de produção é dado pela quantidade de horas de trabalho contratadas (l), multiplicada pela taxa de salários (w), ou seja:

$$CT_i = w_i l_i \quad (5)$$

Seja h a quantidade de horas de trabalho necessárias para a produção de uma unidade do bem. Assumindo uma função de produção muito simples, que associa horas trabalhadas ao produto final (y), tem-se:

$$y_i = \frac{l_i}{h_i} \quad (6)$$

Substituindo na função de custo total, tem-se:

$$CT_i = w_i h_i y_i \quad (7)$$

Com esta função, o custo marginal pode ser aproximado variando a quantidade y em uma unidade ($\Delta y = 1$).

$$\Delta CT_i = w_i h_i \Delta y_i = w_i h_i = c_i \quad (8)$$

Então, o preço de mercado da empresa:

$$p_i = \mu_i w_i h_i \quad (9)$$

Sob as seguintes hipóteses sobre a estrutura econômica do mercado:

- $\mu_i = \mu_f$, ou seja, o *mark-up* é uma decisão da firma (efeito fixo da firma);
- $w_i = w_s$, o salário é uma média setorial (efeito fixo de setor); e
- $h_i = h_i$, as horas são dadas por coeficientes técnicos do insumo (efeito fixo de insumo).

Têm-se em termos logarítmicos a seguinte equação de preço:

$$\dot{p}_{fsi} = \dot{\mu}_f + \dot{w}_s + \dot{h}_i \quad (10)$$

Onde $\dot{x} = \ln x$. Em palavras, um preço observado pode ser decomposto em efeitos associados à firma, ao setor e ao respectivo bem.

2.2.2 Empresa com Múltiplos Insumos de Produção

Nesta seção estende-se a estrutura teórica apresentada na seção anterior, para o caso em que firmas possam utilizar mais do que um insumo de produção.

O custo total de produção é dado pelo valor de mercado da quantidade de insumo empregada por unidade de produto. Seja w^j o custo do insumo j por unidade de produto, e seja h^j a quantidade do insumo j necessária para produzir uma unidade de produto, então, o Custo Total (CT) será:

$$CT_i = \sum_{j=1}^m w_i^j h_i^j y_i \quad (11)$$

E o custo marginal nesse caso será

$$\frac{\partial CT_i}{\partial y_i} = \sum_{j=1}^m w_i^j h_i^j \quad (12)$$

Substituindo no preço do produto

$$p_i = \mu_i \sum_{j=1}^m w_i^j h_i^j = \sum_{j=1}^m \mu_i w_i^j h_i^j = \sum_{j=1}^m e^{\ln \mu_i w_i^j h_i^j} \quad (13)$$

Assumindo a aproximação $e^x \approx 1 + x$

$$p_i = (m_i + \sum_{j=1}^m \ln \mu_i w_i^j h_i^j) \quad (14)$$

E, em termos logarítmicos,

$$\dot{p}_i = \dot{\mu}_i + \sum_{j=1}^m \dot{w}_i^j + \sum_{j=1}^m \dot{h}_i^j + c_i \quad (15)$$

Onde:

$$c_i = m_i + erro \quad (16)$$

Da mesma forma que antes, pode-se assumir que o *mark-up* é um efeito específico à firma, os preços dos insumos específicos ao setor e os coeficientes de utilização, específicos ao insumo, então, mais uma vez

$$\dot{p}_{fsi} = \dot{\mu}_f + \sum_{j=1}^m \dot{w}_s^j + \sum_{j=1}^m \dot{h}_i^j + c \quad (17)$$

Por essa abordagem, o preço de referência de um determinado insumo pode ser obtido a partir da seguinte fórmula:

$$\exp[\bar{p}] = \exp \left[\bar{\mu}_f + \bar{w}_s + \bar{h}_i \right] \quad (18)$$

Novamente, cada um dos componentes à direita representa médias (ou efeitos fixos) para firmas, setores e insumos.

2.2.3 Modelo Empírico para o Estimador Econômico

Das equações derivadas nas seções anteriores, um modelo econométrico de efeitos fixos pode ser empregado para decompor os preços observados no mercado e recuperar os efeitos específicos relevantes para cálculo do preço de referência.

Considerando uma amostra de N bens observados/coletados ao longo de vários períodos, estima-se a seguinte regressão de múltiplos efeitos fixos.

$$\hat{p}_{fsit} = \alpha + \hat{F}_f + \hat{S}_s + \hat{I}_i + \hat{T}_t + \epsilon_{fsit} \quad (19)$$

Ou seja, o logaritmo do preço do bem i , fornecido pela firma f , no setor s , no instante t é decomposto nos respectivos efeitos fixos \hat{F}_f , \hat{S}_s , \hat{I}_i e \hat{T}_t e em uma outra parte não relacionada a estes efeitos, ϵ_{fsit} .

Por conta do grande número de parâmetros a serem estimados, o método de estimação dessa equação se apoia em Guimarães e Portugal (2010), Carneiro, Guimarães e Portugal (2010) e Correia (2016). Esses trabalhos propõem métodos computacionais que permitem lidar com números muito grandes de efeitos fixos.

O Estimador Econômico para o Preço de Mercado do bem i no instante t é composto pelos preços preditos em logaritmo para os N_i fornecedores potenciais desse bem, exponenciados e multiplicados pelo fator de correção do viés² gerado pela transformação logarítmica, $\hat{\gamma}$. Com os preços previstos para cada fornecedor potencial, calcula-se a média, que resulta então no Estimador Econômico, \hat{p}_{sit} , dado pela equação a seguir.

² Sob a hipótese de normalidade do erro, a esperança estatística do fator de correção de tal viés é dada por $e^{\frac{\sigma^2}{2}}$, que sempre é maior do que a unidade. Para uma variância pequena da ordem de 10%, como é o caso do banco de dados da Fipe, a esperança do fator de correção seria de 1,0050. Ou seja, o preço de referência calculado pelo Estimador Econômico não corrigido, estaria em média sendo subestimado em 0,5% (meio por cento). No entanto, para uma dada amostra, ao aplicar o procedimento de estimação desse fator de correção sugerido em Wooldridge (2016) que não necessita da hipótese de normalidade, nada garante que este esteja abaixo de 1.

$$\hat{p}_{sit} = \frac{1}{N_i} \sum_{f=1}^{N_i} \hat{\gamma} e^{\hat{\alpha} + \hat{\beta}_f + \hat{\delta}_s + \hat{\Gamma}_i + \hat{\Gamma}_t} \quad (20)$$

Os detalhes desses procedimentos são apresentados no Anexo 1.

3. DADOS E METODOLOGIA

3.1 DADOS

Nesse estudo são utilizados dados de Preços de Obras Públicas da Fundação Instituto de Pesquisas Econômicas (Fipe). Neste banco de dados observam-se os códigos de produto, fornecedor, e o mês de referência de tais preços. Além disso, existem produtos com periodicidade de coleta mensal, trimestral, semestral ou anual, ou mesmo sem periodicidade definida. São utilizados dados históricos de 30 meses, compreendendo o período de fevereiro de 2017 a julho de 2019.

O banco de dados completo apresenta 12 mil diferentes produtos, em torno de 3 mil fornecedores e 147 setores. Pela natureza do estimador estatístico, sempre que um produto é selecionado para ter seus preços pesquisados, todos os fornecedores cadastrados e conhecidos desse produto são contatados pelos entrevistadores. Esta restrição de ter que contatar todos os fornecedores para formar um preço de referência para o produto pode ser relaxada com a aplicação do estimador econômico³, possibilitando uma estratégia de coleta mais eficiente.

Na Tabela 1 observam-se algumas estatísticas básicas dos 147 setores presentes no banco de dados.

Tabela 1: Descrição dos 147 Setores

Característica dos Setores	Média	D.P.	Mínimo	Máximo
Produtos distintos	81,96	36,16	4	234
Firmas distintas	67,18	34,79	1	207
Produtos-Firmas distintos	296,29	149,83	9	716
Produtos com Preço Observado em 07/2019	21,26	15,81	4	82
Produtos-Firmas Observados em 07/2019	64,17	56,21	8	263

Nota: A tabela reporta estatísticas descritivas dos setores presentes no banco de dados. São reportados a média, o desvio padrão, o mínimo e o máximo. As estatísticas consideram o período de 30 meses e referem-se a produtos distintos observados nos setores, diferentes firmas, e diferentes combinações produtos-firmas, pois nem todas as firmas fornecem todos os produtos de um setor. Além disso, reportam-se as estatísticas de diferentes produtos e combinações de produtos-firmas distintas com ao menos um preço observado no último mês do período, julho de 2019.

Fonte: Fipe, Elaboração Própria.

³ O método permite que cada fornecedor de um determinado produto seja pesquisado em um mês diferente. Ainda assim, tem-se um preço de referência médio, pois têm-se o preço estimado em todos os meses para cada fornecedor que tem ao menos uma observação no intervalo de 30 meses considerado.

A tabela demonstra que em média os setores possuem 81,96 produtos diferentes, 67,18 firmas, que não necessariamente fornecem todos os produtos do setor, e 296,29 combinações de Produtos-Firmas conhecidas. Em julho de 2019 foram observados em média 21,26 produtos em cada setor. No mesmo mês, em média 64,17 Produtos-Firmas tiveram seu preço orçado.

A partir desta base de dados e empregando-se estratégias para seleção de observações é possível testar a performance dos estimadores estatístico e econômico. A próxima subseção descreve as estratégias utilizadas para realizar tal comparação.

3.2 MÉTODO

O estimador estatístico é calculado a partir dos dados observados em um mês específico, para a amostra coletada naquele mês. O estimador econômico, por outro lado, utiliza informações correntes e passadas para calcular o preço de referência do mês. Assim, esses estimadores variam a depender de como as amostras são coletadas ao longo do tempo, bem como em relação ao modo como os choques transitórios ou permanentes de preços são incorporados aos mesmos. Comparamos a performance dos dois estimadores utilizando duas formas alternativas: i) dados simulados e ii) dados reais utilizando preços do Banco de Dados da Fipe.

Para o primeiro caso, de simulação, os dados são gerados supondo parâmetros e distribuições para as variáveis de interesse, isto é, sorteiam-se os efeitos de produto, fornecedor e período, além do termo aleatório não explicado pelo modelo. Com todos os componentes aleatorizados, é possível gerar o preço simulado de cada produto-firma-período. Além dessas, sorteia-se ainda uma variável auxiliar que indica quais preços devem ser “observados”, representando a impossibilidade de se coletar preços em todos os fornecedores do produto. Em seguida calculam-se os Estimadores Econômico e Estatístico para avaliar qual se aproxima mais do Preço de Mercado.

Para o segundo caso, considerando os dados reais do banco de dados da Fipe, aplicam-se dois procedimentos diferentes para comparar os dois estimadores. Antes de discutir cada um desses procedimentos, cabe destacar que uma diferença importante entre o estimador econômico e o estimador estatístico refere-se à forma como cada um deles lida com dados

não observados. O estimador estatístico dá peso apenas para preços efetivamente observados, ao passo que o estimador econômico considera o histórico dos preços disponíveis, mesmo que não observados em um mês específico.

O primeiro procedimento consiste em avaliar a distância entre o Estimador Econômico e o Estatístico. Considerando apenas uma parte dos dados observados para compor a amostra e calcular o Estimador Econômico, avalia-se a distância deste ao Estimador Estatístico, que é computado com toda a informação disponível. Com esse procedimento, pode-se avaliar a distância entre os estimadores fora da amostra, isto é, considerando períodos para os quais os preços de fato observados, não foram utilizados pelo Estimador Econômico.

O segundo procedimento sorteia aleatoriamente uma variável que representa a observação ou não de cada par produto-fornecedor. Supõe-se que a amostra completa, incluindo os sorteados para serem ocultados, representa o Preço de Mercado. Calcula-se cada estimador segundo seu método e compara-se com os supostos Preços de Mercado.

O número de observações por setor em cada mês é em boa medida controlável pela instituição que realiza a pesquisa de preços. Para que choques espúrios nos preços dos produtos pesquisados em dado mês não sejam repassados totalmente ao preço de referência, é preciso ter uma amostra suficientemente grande. Nos testes realizados, sem prejuízo à sua validade, é possível direcionar a análise apenas aos setores em que houve um número satisfatório de observações sorteadas em cada mês.

4. RESULTADOS

Nessa seção são apresentados os diferentes testes realizados para comparar o Estimador Econômico com o Estimador Estatístico, tanto em termos de variabilidade, quanto em termos de consistência, no sentido de se aproximar mais dos supostos parâmetros desejados, ou seja, os Preços de Mercado.

Os exercícios apresentados nas seções 4.1 e 4.3 compartilham um objetivo comum e podem ser diretamente comparados. Ambos exercícios definem o verdadeiro Preço de Mercado, ocultam algumas observações e computam os Estimadores Econômico e Estatístico. Então, ambos estimadores são comparados para checar qual que em média se aproxima mais do Preço de Mercado suposto verdadeiro. A diferença fundamental entre as seções são os dados utilizados. Na primeira os dados são simulados, e por construção os Preços de Mercado são conhecidos. Na última os dados são reais, e se supõe que para cada produto, todos os fornecedores existentes no banco de preços da Fipe constituem todas as firmas de seu mercado.

Por fim, na seção 4.2 é avaliada a distância entre os Estimadores Econômico e Estatístico. Tal seção traz um resultado importante, demonstrando que uma eventual transição da aplicação de um estimador para outro, possui em média uma alteração de magnitude relativamente pequena no nível dos Preços de Referência.

4.1 EXPERIMENTO COM PREÇOS SIMULADOS

No presente exercício de simulação procurou-se calibrar os parâmetros com base na realidade observada no banco de dados do Ipop-Fipe. Geraram-se 100 vezes os dados para um setor hipotético de 100 produtos, com 30 períodos e 100 fornecedores diferentes. Assim como no banco de dados real, nem todas empresas fornecem todos os produtos, e nem todos os fornecedores de cada produto são necessariamente conhecidos pela instituição que pesquisa os preços.

O procedimento adotado para simular o Preço de Mercado e testar os estimadores é descrito da seguinte forma:

1. Cria-se um banco de dados com 100 produtos, 30 períodos e 100 fornecedores, com um efeito médio de setor suposto igual ao número 3;
2. A probabilidade de cada firma fornecer cada um dos Produtos foi estabelecida em 10%, ou seja, em média cada firma fornece 10 produtos do setor, definidos através do sorteio de uma variável indicadora binomial.

$$Fornece_{if} \sim Binomial(1, 0.1),$$

3. A probabilidade definida para um fornecedor nunca ser encontrado pela instituição que pesquisa preços é de 20%, definida pela variável indicadora sorteada pela seguinte distribuição.

$$Fornecedor_f \sim Binomial(1, 0.2),$$

4. Sortearam-se os efeitos médios para cada produto, cada firma e cada período a partir de distribuições normais com média zero.

$$I_i \sim N(0, 2),$$

$$F_f \sim N(0, 0.1),$$

$$T_t \sim N(0.02 t, 0.1)$$

Isto é, supôs-se uma tendência média de aumento de 2% ao mês.⁴

5. Para cada simulação, sorteou-se uma variável no banco de dados que indica a não observação de cada preço, da seguinte maneira:

$$Missing_{ift} \sim Binomial(1, 3/12)$$

Ou seja, cada preço tem a esperança de ser observado 4 vezes a cada 12 meses.

6. Para cada simulação, sorteou-se uma variável com um choque aleatório para cada observação produto-firma-período, da seguinte forma:

$$\varepsilon_{ift} \sim N(0, 0.12)$$

Sendo 12% um número escolhido por se aproximar da média da diferença entre o preço estimado para os produtos-fornecedores e o observado no banco de dados reais utilizados no estudo.

⁴ A não consideração de tal tendência se mostrou indiferente para os resultados obtidos.

7. Desta forma, os preços foram gerados pela seguinte equação:

$$\log(p_{ift}) = 3 + I_i + F_f + T_t + \varepsilon_{ift} \quad (21)$$

E o Preço de Mercado de cada produto em um dado período foi definido pela seguinte fórmula:

$$p_{it} = \frac{1}{N_i} \sum_f e^{\log(p_{ift})} \quad (22)$$

Ou seja, a média do exponencial dos logaritmos dos preços das N_i firmas que fornecem o produto i no período t .

A Tabela 2 apresenta as estatísticas descritivas das 100 simulações realizadas. Para fins descritivos, foram selecionados apenas os 10 primeiros produtos, para os quais se reporta o Preço de Mercado simulado, o Estimador Estatístico e o Estimador Econômico para esse preço, ambos calculados sem as observações sorteadas para serem ocultadas.

Tabela 2: Estatísticas Descritivas das 100 Simulações

Produto	Preço de Mercado	D.P.	Estimador Estatístico	D.P.	Estimador Econômico	D.P.
1	4,03	0,12	4,02	0,37	4,03	0,06
2	1,81	0,06	1,78	0,15	1,81	0,03
3	7,10	0,41	6,96	0,61	7,17	0,18
4	4,05	0,14	4,15	0,39	4,04	0,08
5	39,13	1,57	39,03	4,01	39,27	0,68
6	158,98	5,28	157,87	16,36	158,88	2,72
7	3,26	0,11	3,16	0,27	3,27	0,05
8	81,86	3,36	82,62	9,26	82,21	1,45
9	3,57	0,17	3,61	0,48	3,57	0,09
10	79,23	2,76	80,91	8,32	78,74	1,51

Nota: A tabela apresenta para 10 produtos selecionados, as médias do Preço de Mercado simulado e o respectivo Desvio Padrão para o último período simulado ($t = 30$). Nas colunas seguintes são apresentadas as médias do Estimador Estatístico e do Estimador Econômico, além dos respectivos desvios padrões (D.P.).

Fonte: Fipe, Cálculos Próprios.

Os resultados para as medidas de ajuste calculadas para as 100 simulações combinadas são apresentados na Tabela 3.

Tabela 3: Desvios Absolutos do Estimador Econômico e Estatístico em relação ao Preço de Mercado Simulado

Grupo de Produtos	Obs.	Média	D.P.	Mín	Máx
Todos os produtos					
Estimador Econômico	10.000	3,5%	2,7%	0,0%	23,4%
Estimador Estatístico	8.622	7,9%	6,7%	0,0%	57,2%
Produtos com alguma observação no mês ($n > 1$) mas não todo o mercado ($n < N$)					
Estimador Econômico	8.620	3,3%	2,6%	0,0%	20,8%
Estimador Estatístico	8.620	7,9%	6,7%	0,0%	57,2%
Produtos com nenhuma observação no mês ($n = 0$)					
Estimador Econômico	1.378	4,2%	3,2%	0,0%	23,4%
Estimador Estatístico					
Produtos com muitos fornecedores no Mercado ($N > 5$)					
Estimador Econômico	9.000	3,3%	2,6%	0,0%	17,8%
Estimador Estatístico	7.937	7,8%	6,7%	0,0%	57,2%
Produtos com muitas observações ($n/N > 70\%$) mas não todo o mercado ($n < N$)					
Estimador Econômico	9	2,4%	2,5%	0,1%	7,6%
Estimador Estatístico	9	2,9%	2,2%	0,6%	7,6%
Produtos com poucas observações ($n/N < 30\%$)					
Estimador Econômico	7.794	3,5%	2,7%	0,0%	23,4%
Estimador Estatístico	6.416	8,8%	7,2%	0,0%	57,2%

Nota: Cada cenário desta tabela seleciona um grupo de produtos para reportar as medidas de ajuste. O primeiro cenário compreende todos os produtos. Vê-se que tem menos observações no Estimador Estatístico, pois produtos em que nenhum fornecedor foi observado não possuem Preço de Referência por este estimador. O segundo cenário seleciona todos os produtos em que ao menos um fornecedor teve seu preço selecionado para ser observado, mas não todos fornecedores do mercado. O terceiro apenas os produtos em que nenhum fornecedor teve seu preço observado. O quarto cenário seleciona produtos cujo mercado (N) é composto por muitos fornecedores (> 5), independentemente de quantos destes foram observados. O quinto seleciona produtos em que o instituto de pesquisa teria alcançado a maior parte dos fornecedores existentes ($n / N > 70\%$), mas ao menos um permaneceu desconhecido ($n < N$). O último cenário considera produtos em que o instituto de pesquisa teria conhecimento de uma parte (n) menor do que 30% dos fornecedores do mercado (N).

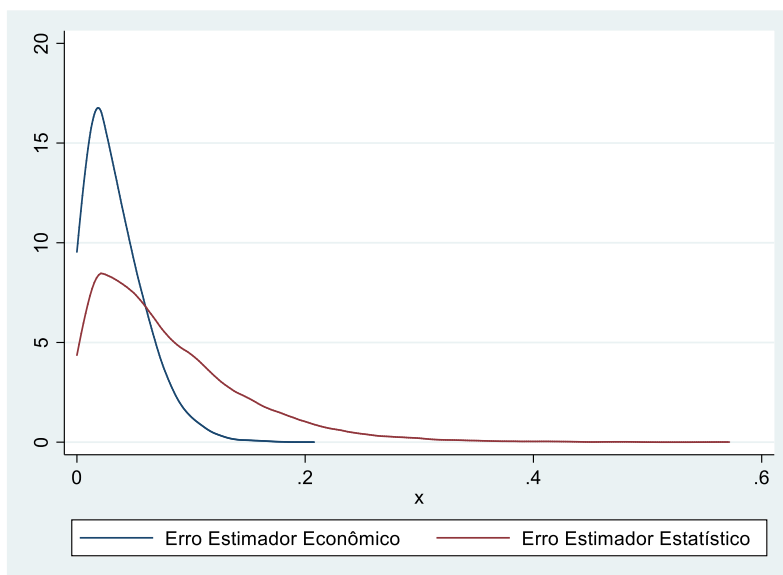
Fonte: Elaboração própria.

Observa-se para o Estimador Econômico um desvio absoluto médio de 3,5% em relação ao verdadeiro Preço de Mercado, enquanto o Estimador Estatístico apresenta desvio de 7,9%. Para os Produtos em que a maior parte ($> 70\%$), mas não todos os fornecedores do mercado foram observados no mês, o Estimador Estatístico reduz seu erro médio para próximo do Econômico, mas ainda assim se mostrou mais elevado (2,9% contra 2,4%). Nas 100 simulações, apenas 9 produtos se enquadraram nesse grupo.

A seguir, na Figura 1 é possível observar as distribuições dos erros relativos absolutos de ambos estimadores em relação ao Preço de Mercado. Para esse gráfico foram selecionados os produtos que na simulação não tiveram todos os fornecedores do mercado pesquisados no mês de interesse, mas tiveram ao menos um. Nota-se que a distribuição

do erro do Estimador Econômico se encontra à esquerda, portanto relativamente mais concentrada em torno do zero.

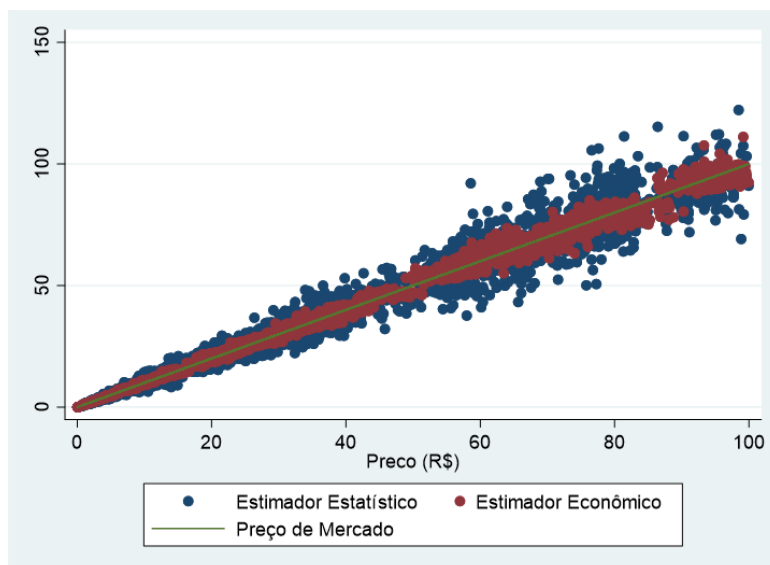
Figura 1: Distribuição do Erro – Estimador Econômico vs. Estimador Estatístico



Fonte: Elaboração própria.

O mesmo resultado pode ser visto na Figura 2, em que cada ponto do gráfico representa o par (Preço de Mercado, Estimador Econômico) em vermelho, (Preço de Mercado, Estimador Estatístico) em azul, e a linha em verde representando o local em que o estimador e o preço verdadeiro coincidiriam.

Figura 2: Preços de Referência - Econômico, Estatístico e o Preço de Mercado



Fonte: Elaboração própria.

Novamente, é possível observar a maior dispersão do Estimador Estatístico em relação ao Econômico.

4.2 DISTÂNCIA ENTRE ESTIMADORES ECONÔMICO E ESTATÍSTICO

Nesse teste investiga-se a distância entre ambos estimadores. Para o cenário que analisa a distância entre o Estimador Econômico e o Estimador Estatístico completamente fora da amostra, consideram-se os produtos em que todas suas observações para o mês de interesse foram ocultadas. Através das observações escondidas conseguimos calcular o que seria o Estimador Estatístico. Por outro lado, mesmo desconsiderando essas observações, através do histórico de dados conseguimos calcular o Estimador Econômico para esses períodos ocultados. As etapas desse procedimento são expostas a seguir.

1. Determinou-se um intervalo para que se observem os preços de todos os produtos. No presente caso, adotou-se intervalo mínimo de uma observação para outra em 12 meses;
2. Para cada produto, selecionou-se um dos meses em que este possui preços observados, e supôs-se que contando deste mês, permanecem considerados apenas os preços observados em meses múltiplos do intervalo definido. Com isso, a cada mês tem-se alguns produtos diferentes com preços observados em cada setor;
3. Com as observações dos meses coincidentes com o intervalo selecionado, calculou-se o Estimador Econômico para todos os períodos;
4. Corrigiu-se os o viés de transformação logarítmica pelo método sugerido em Wooldridge (2016), permitindo que cada setor tivesse um parâmetro de viés diferente;
5. Para reportar as medidas de distância desejadas manteve-se na base apenas o último mês, de julho de 2019, que por ser o mais atual, é o mês de maior interesse no processo. Para computar as medidas de distância dos estimadores, em um dos cenários consideraram-se apenas os produtos sorteados para que as observações de julho fossem escondidas. Dessa forma, obtém-se a medida de distância completamente fora da amostra, pois se tratam dos produtos cujas observações de julho foram descartadas no cálculo do Estimador Econômico.

Após realizar os passos descritos, calculou-se a medida de desvio entre ambos estimadores. Define-se o Desvio Absoluto como o módulo da diferença entre o Preço de Referência gerado pelo Estimador Econômico, e o gerado pelo Estimador Estatístico, supondo a mesma amostra de fornecedores. A fórmula aplicada pode ser representada pela equação a seguir.

$$DA_{econ} = \left| \frac{\widehat{P}_{econ} - \widehat{P}_{estat}}{\widehat{P}_{estat}} \right|$$

Em que \widehat{P}_{econ} representa o Estimador Econômico para um determinado produto em julho de 2019, e \widehat{P}_{estat} representa o Estimador Estatístico. Como resultado do procedimento exposto, obtiveram-se os valores reportados na Tabela 4. Conforme mencionado, nem todos os produtos possuem preços no mês considerado, e dos que possuem, alguns foram escondidos para obter a comparação fora da amostra. Por isso, o número de Produtos considerados em cada medida é consideravelmente menor do que o total de 12 mil produtos constantes na base de dados.

Tabela 4: Distância entre Estimador Econômico e Estatístico

Grupo de Produtos	Núm. Produtos	Média	D.P.
<i>Todos os Produtos</i>			
D.A. (Econômico – Estatístico)	3.022	7,3%	8,5%
<i>Produtos não observados no mês (n = 0)</i>			
D. A. (Econômico – Estatístico)	2.137	8,9%	9,2%
<i>Produtos de Setores com pelo menos 20 observações no mês</i>			
D. A. (Econômico – Estatístico)	1.454	6,4%	7,3%
<i>Produtos não observados de Setores com pelo menos 20 observações no mês</i>			
D. A. (Econômico – Estatístico)	946	7,8%	7,9%

Nota: A tabela reporta o número de produtos em cada cenário do teste, e distância média e o desvio padrão dessa distância entre o Estimador Econômico e o Estatístico. O primeiro cenário seleciona todos os produtos para os quais a Fipe de fato observou preços em julho de 2019, independentemente se foram ou não escondidos para calcular o Estimador Econômico. O segundo cenário compreende apenas os produtos em que a Fipe observou preços em julho de 2019, mas que foram supostos como ausentes para calcular o Estimador Econômico. O terceiro cenário considera o grupo de produtos pertencentes a setores em que após esconder os preços de alguns produtos, sobraram ao menos 20 observações para calcular o Estimador Econômico. O quarto cenário constitui um subgrupo de produtos do terceiro cenário, em que são selecionados apenas os produtos daqueles mesmos setores que foram selecionados para serem escondidos.

Fonte: Fipe, Elaboração própria.

Na Tabela 4 nota-se que a distância entre o Estimador Econômico e o Estimador Estatístico é em média de 8,9%, fora da amostra, isto é, quando os preços deste mês foram utilizados para calcular o Estimador Estatístico, mas não foram utilizados para calcular o

Estimador Econômico do mesmo mês. Tal desvio é relativamente baixo se tomarmos como referência que o desvio padrão de preços dos fornecedores observados de cada produto em julho de 2019 foi, em média, de 20,4%, ou seja, mais do que o dobro da medida de distância média apresentada.

Além disso, é interessante notar que ao selecionar produtos pertencentes a setores que possuíram maior número de observações no mês de julho de 2019, o desvio se mostrou ligeiramente menor, em torno de 7,8%.

A próxima subseção apresenta o último teste desenvolvido, no qual se compara o desempenho de ambos estimadores em se aproximar dos supostos verdadeiros Preços de Mercado.

4.3 EXPERIMENTO COM PREÇOS REAIS

Neste outro teste aplicado ao estimador desenvolvido, utiliza-se novamente o banco de dados do Ipop-Fipe. Agora, supõe-se que o preço médio de todos os fornecedores observados para um produto em cada mês constitui o Preço de Mercado. Ao sortear aleatoriamente algumas observações para serem desconsideradas, calculam-se os Estimadores Econômico e o Estatístico, e compara-se com o suposto parâmetro verdadeiro. O procedimento é detalhado a seguir:

1. Com a base histórica de preços de fevereiro de 2017 a julho de 2019, sorteou-se uma variável aleatória Binomial (1, 0.3), ou seja, com 30% de chance de sucesso, que no caso significa que o preço Insumo-Fornecedor-Período será escondido na estimação;
2. Aplicou-se então o Estimador Econômico, e o Estimador Estatístico às observações não escondidas;
3. Com a hipótese de que o Preço Médio considerando todos os preços de cada período (ou seja, também os 30% escondidos), é o Preço de Mercado, realizaram-se comparativos entre os dois estimadores para o mês de interesse, com as fórmulas similares às apresentadas na seção anterior, reportadas a seguir:

$$DA_{econ} = \left| \frac{\widehat{P}_{econ} - P}{P} \right|$$

$$DA_{estat} = \left| \frac{\widehat{P}_{estat} - P}{P} \right|$$

Em que P constitui o suposto verdadeiro Preço de Mercado.

A seguir, a Tabela 5 reporta os resultados obtidos para diferentes grupos de produtos. Em linhas gerais nota-se que o erro do Estimador Econômico é menor do que o Estimador Estatístico, sendo maior apenas quando se consideram todos os produtos. Neste último caso, estão incluídos os produtos em que se observam todos os fornecedores cadastrados no mês, caso em que o erro do estatístico é zero, por construção. Além disso, tal fato é esperado, pois quando a amostra observada no mês vai aumentando, o Estimador Estatístico converge para a própria definição suposta para o Preço de Mercado.

**Tabela 5: Desvio Absoluto em relação ao Preço de Mercado:
Estimador Econômico vs. Estatístico**

Estimador	Obs.	Média	Mín	Máx
<i>Todos os produtos</i>				
Econômico	3.111	5,2%	0,0%	114%
Estatístico	2.775	4,1%	0,0%	58,2%
<i>Produtos com ao menos uma observação no mês (n > 1) mas não todas (n < N)</i>				
Econômico	1.457	4,5%	0,0%	57,3%
Estatístico	1.457	7,9%	0,0%	58,2%
<i>Produtos com nenhuma observação no mês (n = 0)</i>				
Econômico	336	8,4%	0,0%	114%
Estatístico	ND	ND	ND	ND
<i>Produtos com muitos fornecedores no Mercado (N > 5)</i>				
Econômico	380	3,2%	0,0%	19,9%
Estatístico	380	4,1%	0,0%	25,8%
<i>Produtos com muitas observações (n/N > 70%) mas não todo o mercado (n < N)</i>				
Econômico	410	3,9%	0,0%	25,1%
Estatístico	410	4,3%	0,0%	20,6%
<i>Produtos com poucas observações (n/N < 30%)</i>				
Econômico	394	7,9%	0,0%	114%
Estatístico	58	13,7%	0,0%	47,5%

Nota: Cada cenário desta tabela seleciona um grupo de produtos para reportar as medidas de ajuste. O primeiro cenário compreende todos os produtos. Vê-se que tem menos observações no Estimador Estatístico, pois produtos em que nenhum fornecedor foi observado não possuem Preço de Referência por este estimador. O segundo cenário seleciona todos os produtos em que ao menos um fornecedor teve seu preço selecionado para ser observado, mas não todos fornecedores do mercado. O terceiro apenas os produtos em que nenhum fornecedor teve seu preço observado. O quarto cenário seleciona produtos cujo suposto mercado (N) é composto por muitos fornecedores (> 5) independentemente de quantos destes foram observados. O quinto seleciona produtos em que o instituto de pesquisa alcançou a maior parte dos fornecedores supostos como existentes (n / N > 70%) mas ao menos um permaneceu desconhecido (n < N). O último cenário considera produtos em que o instituto de pesquisa alcançou uma parte (n) menor do que 30% dos fornecedores do suposto mercado (N).

Fonte: Fipe. Cálculos próprios.

Caso se deseje considerar o ajuste fora da amostra, o Estimador Estatístico para o Preço de Mercado não está definido, mas o Estimador Econômico existe e apresenta desvio médio de 8,4%, o que, conforme mencionado em seção anterior, é relativamente baixo ao se comparar com o Desvio Padrão médio de 20,4% dos preços dos fornecedores de um mesmo produto em um mesmo mês.

Ao serem selecionados os produtos para os quais se observa mais de 70% dos fornecedores do mercado, o Estimador Econômico ainda apresenta um ajuste melhor ao Preço de Mercado do que o Estimador Estatístico. É interessante notar que quando observamos menos do que 30% do mercado, ambos estimadores pioram o ajuste, o que é esperado, pois se está perdendo muita informação, mas o Estimador Econômico mostra uma distância muito menor do Preço de Mercado, de 7,9%, do que o Estimador Estatístico, que apresenta desvio médio de 13,7%.

Em suma, a única situação em que o Estimador Estatístico se ajustou mais, em média, ao “Preço de Mercado”, ocorre quando selecionamos todos os produtos, grupo que inclui os casos em que observamos todos dos fornecedores do mercado no mês de interesse (julho 2019). Este cenário se mostra improvável, pois mesmo que se queira obter os preços de todos os fornecedores, atritos de campo impedirão, e seria altamente custoso, pois deveriam ser entrevistados todos os fornecedores em todos os meses de todos os produtos.

Na Tabela 6, a seguir, faz-se um apanhado dos principais resultados obtidos nos diferentes testes deste artigo. Em resumo, fica evidente que para o caso de pequenas amostras, e quando não é possível alcançar todos os fornecedores do mercado de um produto, o Estimador Econômico se mostrou uma alternativa melhor e mais robusta do que a média dos preços observados, aqui denominado de Estimador Estatístico.

Tabela 6: Resumo de Resultados

Seção de Testes	Medida	Estimador Econômico	Estimador Estatístico
Distância entre os Estimadores	D.A. Médio ao Preço Médio Observado	7,3%	0
Experimento Preços Reais	D.A. Médio ao Preço de Mercado	4,5%	7,9%
Experimento Preços Simulados	D.A. Médio ao Preço de Mercado	3,3%	7,9%

Fonte: Fipe. Cálculos próprios.

Sobre os três tipos de testes aplicados, pode-se sumarizar os principais resultados da seguinte maneira:

- Ajuste ao Estimador Estatístico (preço médio pesquisado): Preços de Referência obtidos pelo Estimador Econômico são próximos aos obtidos pelo Estimador Estatístico, mesmo quando são realizadas comparações fora da amostra;
- Experimento com Preços Reais: Mesmo supondo que a Coleta alcançasse todo o mercado, ao esconder parte dele, o Estimador Econômico se aproxima mais do Preço de Mercado do que o Estimador Estatístico;
- Experimento com Preços Simulados: Em um ambiente simulado em que conhecemos todos os membros do mercado e seus preços, ao não observarmos parte dele, o Estimador Econômico se aproxima mais do Preço de Mercado verdadeiro do que o Estimador Estatístico.

5. CONCLUSÃO

Através dos exercícios e medidas apresentadas, demonstrou-se que o Estimador Econômico para Preço de Mercado aplicado aos preços pesquisados em entrevistas aos fornecedores diminui variações não justificáveis economicamente, constituindo assim um método alternativo e mais robusto para construção de um Preço de Referência ao Setor Público.

Em linhas gerais, demonstrou-se que a adoção de um modelo econométrico que toma vantagem do conteúdo informacional de um grande banco de preços permite a redução do número de visitas aos fornecedores ao longo do ano, mantendo preços de referência mais estáveis e atualizados do que o Estimador Estatístico, e, portanto, considerados mais robustos para os usuários do Setor Público. Além disso, nos testes realizados em que se supõe o verdadeiro Preço de Mercado, o Estimador Econômico apresentado performou melhor ao se aproximar mais, em média, do que o Estimador Estatístico.

Uma característica da coleta de campo, e que pode facilmente ser vista nos dados, é sua extrema variabilidade, com aumentos e quedas sucessivas e de várias ordens. Algumas podem realmente refletir o preço de mercado naquele fornecedor (promoção), mas tal medida pode não ser adequada para refletir o preço de mercado que será praticado na obra pública, ou para embasar reajustes de contratos públicos. Outras variações são puros erros de medida, como ocorre quando se muda o informante habitual, o coletor de preços ou qualquer outro erro de anotação do questionário, digitação etc. Não é incomum que esses erros na coleta sejam corrigidos em meses posteriores, introduzindo na série de preços uma variação espúria: um aumento que seria seguido por queda, ou vice-versa, quando na verdade o “preço de mercado” não teve qualquer variação.

O Estimador Econômico nada mais é do que uma outra abordagem, igualmente estatística, para se determinar o que é um preço de mercado a partir de uma amostra de preços. A grande diferença é que a amostra não é definida apenas por algumas poucas observações em um determinado período, mas é composta por observações atuais e pretéritas do preço do insumo, bem como de características do mercado e do fornecedor.

O método - que deve ser visto como uma família potencial de modelos estatísticos – utiliza um conjunto de informação maior para obter o preço de mercado, sendo, por isso, mais robusto. De fato, o modelo atual é menos sensível a variações contemporâneas, mas essa é uma de suas principais vantagens. A depender da situação, admitir uma variação de preços quando ela não ocorreu, pode ser mais danoso do que suavizá-la quando verdadeira. Sobre esse ponto, o procedimento tradicional, ou o Estimador Estatístico, previa que uma variação de preços, que o analista entendia não ser verdadeira, pudesse ser incorporada apenas em meses posteriores, caso se confirmasse nas coletas seguintes. O que o modelo atual faz é tornar esse procedimento estatisticamente justificável, sem depender do julgamento humano, que pode ser falho.

6. REFERÊNCIAS

- ABOWD, J. M. et al. (2002) - Computing person and firm effects using linked longitudinal employer-employee data. Center for Economic Studies, US Census Bureau.
- CARNEIRO, A., P. GUIMARÃES, P., PORTUGAL, P. (2010). Real wages and the business cycle: Accounting for worker, firm, and job heterogeneity. Unpublished manuscript.
- CAVALLO, A., RIGOBON, R. (2016) The billion prices project: Using online prices for measurement and research. *Journal of Economic Perspectives* 30, nº 2 (2016): 151-78.
- CORREIA, S., (2016). REGHDFE: Estimating linear models with multi-way fixed effects. Stata Users Group. disponível em <https://ideas.repec.org/s/boc/scon16.html>
- GUIMARAES, P., PORTUGAL, P. (2010) A simple feasible procedure to fit models with high-dimensional fixed effects. *The Stata Journal*, v. 10, n. 4, p. 628-649.
- SANTOS, F. B. (2015) Preço de Referência em Compras Públicas (Ênfase em Medicamentos) – Cartilhas, Manuais e Tutoriais - TRIBUNAL DE CONTAS DA UNIÃO (TCU) – disponível em: <https://portal.tcu.gov.br/biblioteca-digital/preco-de-referencia-em-compras-publicas-enfase-em-medicamentos.htm>
- SILVA FILHO, L. O., LIMA, M. C., MACIEL R. G. (2010) Efeito barganha e cotação: fenômenos que permitem a ocorrência de superfaturamento com preços inferiores às referências oficiais. *Revista do TCU*. disponível em <https://revista.tcu.gov.br/ojs/index.php/RTCU/article/view/188>
- SINAPI (2017) Sistema nacional de pesquisa de custos e índices da construção civil: Métodos de Cálculo/IBGE, Coordenação de Índices de Preços. – Rio de Janeiro: IBGE, 2017.

WOOLDRIDGE, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.

WOOLDRIDGE, J. M., (2016) – Introductory Econometrics: A modern approach – Ed. Cengage Learning 6th Ed.

ANEXO 1: MODELO ECONÔMETRICO

Cada preço é decomposto em um efeito médio de insumo, um efeito médio de fornecedor, e um efeito médio do mês. Adaptamos para este fim a metodologia iniciada por Abowd (2002) e que ainda segue em evolução (Correia, 2016), que lida com a presença de grupos grandes de efeitos fixos. No presente caso emprega-se uma base de dados com mais de 10 mil insumos, e com 4 grupos de efeitos, Setor, Insumo, Empresa e Tempo.

Com as seguintes hipóteses, podemos estipular um modelo econômico para os preços de referência:

- Produtos similares estão sujeitos a componentes temporais comuns;
- Fornecedores diferem enquanto a qualidade e nível de preço em cada setor de mercado;
- Cada produto possui um preço de equilíbrio de médio prazo, que só é deslocado em caso de choques tecnológicos ou demanda permanentes.

O método aplicado para gerar os preços de referência pode ser resumido da seguinte maneira: dividem-se os insumos em grupos de produtos com algum grau de similaridade, chamados de Grupos de Inferência; com algumas observações de cada produto, de cada fornecedor e em cada mês, estima-se a seguinte equação:

$$\log(p_{fsit}) = \alpha + S_s + I_i + F_f + T_t + \varepsilon_{ift} \quad (1)$$

Onde s indica o setor, i indica produto, f indica a firma e t indica o mês de referência da observação. Este modelo decompõe cada preço em um efeito médio do setor (S), um de produto (I), um de empresa/fornecedor (F), um de tempo (T), e em um choque aleatório (ε) não observado.

Em linguagem econométrica, este modelo é chamado de Modelo de Efeitos Fixos⁵, e seria estimado incorporando variáveis indicadoras (*dummies*) para cada setor, produto,

⁵ Embora os Efeitos Fixos sejam atualizados a cada mês, optou-se por manter o nome tradicionalmente usado na literatura de Econometria de Painel (Wooldridge, 2010).

fornecedor e período. Por razões computacionais⁶, tal procedimento pode ser contornado por algoritmos mais eficientes, chamados de *High Dimensional Fixed Effects* (Abowd et. al., 2002, Guimarães e Portugal, 2010 e Correia, 2016), que numericamente são equivalentes ao resultado de Mínimos Quadrados incorporando variáveis indicadoras.

Uma vez estimada essa equação, é possível selecionar qual amostra de fornecedores quer-se considerar em cada produto e cada mês⁷, e tomar a média dos preços estimados dos fornecedores considerados no mês de referência.

A cada novo período, estima-se este modelo com uma janela temporal dos últimos 30 meses. Isto é, as informações mais antigas passam a não ter peso nos efeitos estimados de produto, fornecedor e período, e por consequência no preço de referência. Desta forma, caso o produto tenha uma tendência diferente da do seu setor, conforme novas observações deste produto entrem na base de dados com o passar dos períodos, o efeito fixo de produto absorverá gradualmente tal tendência. O mesmo vale para tendências no efeito específico de firma.

Sendo o modelo estimado em logaritmo, para reconstruir o preço médio consideramos o exponencial de cada preço estimado antes de realizar a média. Tal operação gera um viés de estimação (Wooldridge, 2016 pg. 190-192), que pode ser resolvido com o procedimento de regredir o preço contra o preço previsto sem a correção e sem intercepto. Isto é, com o preço estimado por (1), realiza-se a seguinte regressão:

$$p_{ift} = \gamma \hat{p}_{ift} + \xi_{ift} \quad (2)$$

Em que:

$$\hat{p}_{ift} = \exp(\hat{\alpha} + \hat{S}_s + \hat{I}_i + \hat{F}_f + \hat{T}_t) \quad (3)$$

⁶ A inclusão de uma variável para cada produto, cada fornecedor e cada período pode trazer problemas por falta de capacidade computacional. Por isso, o método alternativo de solução numérica equivalente se mostra relevante ao procedimento.

⁷ Usualmente se está interessado no mês mais recente da janela temporal de 30 meses. No mês seguinte entrarão observações novas e serão desconsideradas as observações do mês mais antigo.

Estimando γ por (2) obtém-se o preço estimado corrigido:

$$\widehat{p}_{ift} = \hat{\gamma} \hat{p}_{ift} \quad (4)$$

Na seção de testes, tanto com preços simulados quanto com preços reais, sempre são reportadas as medidas de ajuste para os preços corrigidos por este procedimento.