



# ECONOMIA APLICADA

BRAZILIAN JOURNAL OF APPLIED ECONOMICS

Vol. 19 – nº 1

Janeiro – Março 2015

## ARTIGOS

**Índice de preços hedônicos para imóveis: uma análise para o município de Belo Horizonte**  
Luiz Andrés Ribeiro Paixão

**Bolsa Família ou desempenho da economia? Determinantes da reeleição de Lula em 2006**  
Maurício Canêdo-Pinheiro

**Análise da eficiência técnica de propriedades citrícolas do Estado de São Paulo**  
Felipe Clemente, Marília Fernandes Maciel Gomes, Viviani Silva Lório

**Reavaliando a relação entre produtividade e exportações através de modelo de correção de erros não linear: evidências para Estados Unidos, Canadá, Japão e Alemanha**  
João Paulo Martin Faleiros, Denisard Cneio de Oliveira Alves

**Impactos da alteração de perfil do consumo das famílias sobre o emprego por ocupação e nível de qualificação: aplicação de um modelo multissetorial para o Brasil**  
Adelar Fochezatto, Carlos Eduardo Lobo e Silva

**Efeitos dos pisos salariais estaduais sobre o mercado de trabalho: uma nova abordagem empírica**  
Carlos Henrique Corseuil, Miguel Foguel, Marcos Hecksher

**Endividamento Agrícola: Quão Comprometidos são os produtores do polo Petrolina-Juazeiro frente a suas dívidas?**  
Emanuel de Souza Barros, Leonardo Ferraz Xavier, Danielle Tavares Pessoa, Tiago Farias Sobel

ECONOMIA FEA-RP/USP



# ECONOMIA APLICADA

Vol. 19 – nº 1 Janeiro – Março 2015

PAPERS

**The hedonic pricing method used as a price changing measurement tool of real estate: an application for the real estate market in Belo Horizonte**

Luiz Andrés Ribeiro Paixão

**Bolsa Família or Economic Performance? Determinants of Lula's Re-election in 2006**

Mauricio Canêdo-Pinheiro

**Technical Efficiency Analysis of citrus properties from São Paulo**  
Felippe Clemente, Marília Fernandes Maciel Gomes, Viviani Silva Lório

**Revaluating the relationship between productivity and exports through nonlinear error correction model: empirical evidences for United States, Canada, Japan and Germany**

João Paulo Martin Faleiros, Denisard Cneio de Oliveira Alves

**Impacts of changing consumption pattern on employment by occupation and level of qualification: application of a multisectoral model for Brazil**

Adelar Fochezatto, Carlos Eduardo Lobo e Silva

**The effects of local wage floors for labor market outcomes: a new empirical approach**

Carlos Henrique Corseuil, Miguel Foguel, Marcos Hecksher

**Agricultural debt: How committed are the producers in the Petrolina-Juazeiro region to their debt burdens?**

Emanuel de Souza Barros, Leonardo Ferraz Xavier, Danielle Tavares Pessoa, Tiago Farias Sobel

ISSN 1413-8050



APOIO:



Ministério  
da Educação

Ministério da  
Ciência e Tecnologia





# ECONOMIA APLICADA

Vol. 19 – nº 1

Janeiro – Março 2015

## Editor

Reynaldo Fernandes

## Editores Associados

Luiz Guilherme Scorzafave	Sergio Naruhiko Sakurai
Angelo Costa Gurgel	Eliezer Martins Diniz
Rodrigo De Losso	Irineu E. de Carvalho Filho
Cláudio Ribeiro de Lucinda	Walter Belluzzo Junior

## Conselho Editorial

<b>Affonso Celso Pastore</b> USP	<b>Antônio Barros de Castro</b> UFRJ
<b>Cássio F. Camargo Rolim</b> UFPR	<b>Cláudio Monteiro Considera</b> UFF
<b>Clélio Campolina Diniz</b> CEDEPLAR	<b>Denisard C. de Oliveira Alves</b> USP
<b>Eleutério F. S. Prado</b> USP	<b>Fernando de Holanda Barbosa</b> FGV-UFF
<b>Francisco Anuatti Neto</b> USP/RP	<b>Geoffrey J. D. Hewings</b> University of Illinois
<b>Geraldo Sant'ana de Camargo Barros</b> ESALQ/USP	<b>Gustavo Maia Gomes</b> IPEA
<b>José Marcelino da Costa</b> NAEA/PA	<b>José A. Scheinkman</b> Princeton University
<b>Marcelo Portugal</b> UFRGS	<b>Maria José Willumsen</b> Florida International University
<b>Márcio Gomes Pinto Garcia</b> PUC/ RJ	<b>Mário Luiz Possas</b> UFRJ
<b>Paulo César Coutinho</b> UnB	<b>Paulo Nogueira Batista Júnior</b> FGV/SP
<b>Pierre Perron</b> Boston University	<b>Pedro Cesar Dutra Fonseca</b> UFRGS
<b>Ricardo R. Araújo Lima</b> UnB	<b>Robert E. Evenson</b> Yale University
<b>Roberto Smith</b> UFCE	<b>Rodolfo Hoffmann</b> ESALQ/USP
<b>Rogério Studart</b> UFRJ	<b>Russell E. Smith</b> Washburn University
<b>Sérgio Werlang</b> FGV/RJ	<b>Tomás Málaga</b> FGV/SP
<b>Victor Bulmer-Thomas</b> University of London	<b>Werner Baer</b> University of Illinois
<b>Wilson Suzigan</b> Unicamp	



A Revista ECONOMIA APLICADA é uma publicação trimestral do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo. Os conceitos emitidos são de responsabilidade exclusiva do(s) autor(es). A simples remessa do trabalho implica que o(s) autor(es) concordam que, em caso de aceitação para publicação, a ECONOMIA APLICADA (BRAZILIAN JOURNAL OF APPLIED ECONOMICS) passa a ter os direitos autorais para a veiculação dos artigos, tanto em formato impresso como eletrônico a eles referentes, os quais se tornarão propriedade exclusiva da ECONOMIA APLICADA (BRAZILIAN JOURNAL OF APPLIED ECONOMICS). É permitida a reprodução total ou parcial dos artigos desta revista, desde que seja explicitamente citada a fonte completa. Revista financiada com recursos do Programa de Apoio a Publicações Científicas do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, entidade governamental brasileira promotora do desenvolvimento científico e tecnológico. Esta revista está sendo indexada por Journal of Economic Literature, versões *eletronic on line* e CD-ROM e Scielo Brasil.

Secretaria e Divulgação: Juliana de Freitas Vitória      Capa: Christof Gunkel  
Revisão linguísticas (Português e Inglês): Iulo Feliciano Afonso (Solution Translations  
Eireli — Epp)  
Diagramação: Francisco Nunes Navarro & Bruna Carolina Garcia

## Endereço

Revista Economia Aplicada, Departamento de Economia, FEA-PR/USP  
Avenida do Bandeirantes, 3900, Ribeirão Preto, SP.  
CEP 14040-905  
Telefone: (0xx16) 3315-3910  
email:revecap@usp.br  
site: <http://www.usp.br/revecap>

## Assinaturas

tipo	Brasil	Exterior*
Individual	R\$50,00	R\$ 90,00
Institucional	R\$70,00	R\$110,00

\*inclui custos de postagem

### FICHA CATALOGRÁFICA

Elaborada pela Seção de Publicações e Divulgação do SBD/FEA/USP.

Economia Aplicada / Departamento de Economia, Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo – v. 1, n. 1 – Ribeirão Preto: FEA-RP, 1997-

Trimestral

Publicado anteriormente sob responsabilidade de FEA/USP, FIPE e FEA=RP  
ISSN 1413-8050

1. Economia I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto. Departamento de Economia.

CDD 22.ed – 330.

# ECONOMIA APLICADA

Vol. 19 – nº 1

Janeiro – Março 2015

---

## ARTIGOS

- 5 **Índice de preços hedônicos para imóveis: uma análise para o município de Belo Horizonte**  
Luiz Andrés Ribeiro Paixão
- 31 **Bolsa Família ou desempenho da economia? Determinantes da reeleição de Lula em 2006**  
Mauricio Canêdo-Pinheiro
- 63 **Análise da eficiência técnica de propriedades citrícolas do Estado de São Paulo**  
Felippe Clemente , Marília Fernandes Maciel Gomes , Viviani Silva Lírio
- 81 **Reavaliando a relação entre produtividade e exportações através de modelo de correção de erros não linear: evidências para Estados Unidos, Canadá, Japão e Alemanha**  
João Paulo Martin Faleiros , Denisard Cneio de Oliveira Alves
- 109 **Impactos da alteração de perfil do consumo das famílias sobre o emprego por ocupação e nível de qualificação: aplicação de um modelo multissetorial para o Brasil**  
Adelar Fochezatto , Carlos Eduardo Lobo e Silva
- 131 **Efeitos dos pisos salariais estaduais sobre o mercado de trabalho: uma nova abordagem empírica**  
Carlos Henrique Corseuil , Miguel Foguel , Marcos Hecksher
- 171 **Endividamento Agrícola: Quão Comprometidos são os produtores do polo Petrolina-Juazeiro frente a suas dívidas?**  
Emanoel de Souza Barros , Leonardo Ferraz Xavier , Danielle Tavares Pessoa , Tiago Farias Sobel



# ÍNDICE DE PREÇOS HEDÔNICOS PARA IMÓVEIS: UMA ANÁLISE PARA O MUNICÍPIO DE BELO HORIZONTE

LUIZ ANDRÉS RIBEIRO PAIXÃO \*

## Resumo

Este trabalho tem como objetivo mensurar a valorização dos preços dos diversos tipos de imóveis para o período entre 1995 e 2003. Para isso foi utilizada uma amostra contendo valores de transações com apartamentos, casas, barracos, lojas, salas, galpões e terrenos na cidade de Belo Horizonte. Foi utilizado o método dos preços hedônicos, pois este calcula índices de preços controlando pelas características dos bens. Como resultado temos que a valorização dos imóveis foi pequena no período, em muitos casos abaixo da inflação, devido ao fraco desempenho da economia, da alta taxa de juros e do pequeno volume de crédito habitacional.

**Palavras-chave:** Índice de preços para bens imóveis; Modelo de preços hedônicos; Economia.

## Abstract

This paper focuses on measuring the price appreciation of several real estate types between 1995 and 2003. We used as sample transaction values of apartments, houses, huts, shops, halls, warehouses and land in Belo Horizonte. The hedonic pricing method was used once its basic premise is that the price of a marketed good is related to its characteristics. As a result, we could realize that the properties valuation was small in the period, in many cases below the inflation rate, due to the poor economy, high interest rates, and housing credit small volume.

**Keywords:** Real estate price indexes; Hedonic price model; Urban economics; Belo Horizonte.

**JEL classification:** C43, E31, R31.

**DOI:** <http://dx.doi.org/10.1590/1413-8050/ea36708>

---

\* Aluno do Doutorado em Economia do PPGE/IE/UFRJ e Analista do IBGE. E-mail: [lui-pai@yahoo.com](mailto:lui-pai@yahoo.com)

## 1 Introdução

Índice de preços que mensure a variação de preço dos bens imóveis é um indicador importante, tanto em análises macroeconômicas, quanto nos estudos e na prática de planejamento urbano. Todavia, mensurar a variação do preço dos imóveis não é trivial. O bem imóvel é heterogêneo quanto às características que compõem cada unidade; desse modo, índices calculados a partir de valores médios ou medianos podem apresentar viés. Somado a isso, as vendas de cada unidade se dão de forma infrequente no tempo, não sendo possível acompanhar o comportamento de uma cesta predefinida de imóveis ao longo do tempo.

O modelo de preços hedônicos tem sido utilizado como metodologia para se gerar índices de preços para bens imóveis. Por esse modelo, o preço do bem é uma função dos atributos que o compõe. Cada atributo, portanto, teria um preço — embutido no preço final do bem (“preço sombra”) — e o preço do imóvel seria o somatório do preço sombra de cada um de seus atributos. Desse modo, a utilização do modelo de preços hedônicos é uma forma de se obter índices de preço para imóveis controlado pelas diferenças nos atributos de cada observação. Em 2006, a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) e o Fundo Monetário Internacional (FMI) realizaram o *Workshop on Real Estate Price Indexes*, onde foram apresentadas as experiências de vários países na construção de índice de preços para imóveis. Em um balanço das apresentações, Diewert (2009) constatou que o modelo de preços hedônicos é o mais utilizado e, segundo esse autor, é o mais recomendando para a construção dos índices.

Na academia, a aplicação do modelo de preços hedônicos para o mercado de imóveis é diversa, sendo a construção de índice de preços um subitem dessa literatura. Como exemplo da diversidade de temas incorporados pela literatura de preços hedônicos tem-se: Freeman (1979), que utilizou o modelo de preços hedônicos para mensurar o impacto da qualidade do ar no preço dos imóveis; Clark (2006), que teve como objetivo testar o impacto da poluição sonora no mercado imobiliário; Tita et al. (2006) utilizaram o modelo para estimar a desvalorização dos imóveis decorrentes da criminalidade urbana; Gibbons & Machin (2003) tiveram como foco o impacto da qualidade das escolas no valor dos imóveis; Sheppard (2010), por sua vez, testou o impacto das amenidades culturais no valor das propriedades urbanas.

Como exemplo de aplicação do modelo de preços hedônicos na construção de índice de preços para imóveis, tem-se: Gatzlaff & Ling (1994), que estimaram índices para a Região Metropolitana de Miami-EUA entre 1971-1991 utilizando variadas especificações do modelo de preços hedônicos, assim como o modelo das vendas repetidas, não encontrando diferenças significativas nos índices gerados pelas diversas metodologias. Hoesli et al. (1997) utilizaram o modelo de preços hedônicos para o mercado imobiliário de Genebra-Suíça entre 1970-1999. Como resultado, os autores apontam que o preço dos imóveis acompanhou a inflação nos anos de 1970; na década de 1980, houve um boom nesse mercado, contemporâneo ao boom no mercado americano; já na década de 1990 houve um crash em decorrência do aumento na taxa de juros das hipotecas.

Maurer et al. (2004) aplicaram um modelo de preços hedônicos para o mercado imobiliário de Paris-França entre 1990 e 1999. Os autores verificaram uma baixa correlação entre a variação do preço dos imóveis e a variação do

valor das ações e títulos, concluindo que a aquisição de imóveis, ou de títulos lastreados no preço dos imóveis, é importante na diversificação do portfólio dos investidores. Li et al. (2006) testaram varias especificações do modelo de preços hedônicos na geração de índice de preços para Ottawa-Canada, concluindo que os resultados foram praticamente idênticos nas diversas especificações.

No Brasil, o modelo de preços hedônicos também tem sido aplicado para se mensurar diversos aspectos do mercado imobiliário. Alguns estudos focam em um conjunto de variáveis ambientais, como Hermann & Haddad (2005) para São Paulo, Albuquerque et al. (2007), para Recife, e Batalhone et al. (2002), para Brasília. Outros estudos focam em aspectos mais específicos: Teixeira & Serra (2006) testam o impacto da criminalidade urbana para a cidade de Curitiba; Amrein (2010) teve como foco a qualidade das escolas no município de São Paulo; Aguirre & Faria (1997) e Faria (2008) utilizaram o modelo para avaliar políticas públicas.

A construção de índice de preços tem recebido menor atenção nos estudos brasileiros. A literatura sobre o tema ainda é pequena, resumindo-se aos trabalhos de González (1997b), Rozenbaum & Macedo-Soares (2007) e Rozenbaum (2009). González (1997b) mensurou a variação mensal dos aluguéis de apartamentos em Porto Alegre entre janeiro de 1993 e abril de 1997, utilizando o modelo de preços hedônicos. Como resultado, nos primeiros meses do Plano Real (até metade de 1995) os preços subiram significativamente; posteriormente houve queda nos preços, que voltaram ao patamar do início do plano.

Rozenbaum & Macedo-Soares (2007) utilizaram o modelo de preços hedônicos para mensurar a variação mensal do preço de lançamentos no bairro da Barra da Tijuca, Rio de Janeiro, entre 2002 e 2005. Houve desvalorização em 2002 e em 2003 devido ao pequeno dinamismo da economia e o alto patamar da taxa de juros. Já em 2005, houve uma pequena valorização real.

Rozenbaum (2009) calculou um índice trimestral de preço para apartamentos no município do Rio de Janeiro, entre 1999 e 2007, por meio de modelagem hedônica, utilizando dados do Imposto de Transmissão Imobiliária Inte-Vivos (ITBI). Pelos índices calculados, o valor do preço dos apartamentos no Rio de Janeiro acompanhou a inflação durante o período, estando em um patamar bem abaixo da taxa básica de juros da economia (taxa SELIC).

O objetivo deste artigo é mensurar a variação de preço de imóveis para Belo Horizonte, entre 1995 e 2003, a partir do modelo de preços hedônicos, utilizando a base de dados do ITBI de Belo Horizonte. Serão gerados índices anuais para cada tipo de imóvel existente na base do ITBI: apartamento, casa, barracão, terreno, sala, loja e galpão. Este trabalho justifica-se pelas seguintes razões:

- a) A discussão nacional, como apontado acima, é pequena e, portanto, este artigo contribui para o crescimento dessa literatura.
- b) Os artigos existentes na literatura nacional mensuram índices de preços para apartamentos. Neste artigo, estendemos a construção de índice para outros tipos de imóveis, englobando, além de apartamentos, casas, barracões, terrenos, salas, lojas e galpões, construindo índices separados para cada tipo de imóvel.

- c) Foco no município de Belo Horizonte. As contribuições existentes na literatura nacional têm como foco Porto Alegre (González 1997b) e Rio de Janeiro (Rozenbaum & Macedo-Soares 2007, Rozenbaum 2009). Existem estudos utilizando a metodologia hedônica para Belo Horizonte, porém nenhum teve como objetivo a construção de índice de preços. Dos trabalhos existentes para Belo Horizonte, podemos citar: Aguirre & Macedo (1996), que utilizaram dados do mercado imobiliário para estimar os preços sombras de diversas características físicas dos apartamentos; Paixão (2010), Miranda & Amaral (2009) e Furtado (2007), que tiveram como objetivo estimar o impacto de uma série de variáveis de localização no preço de apartamentos; Macedo (1998), que buscou aplicar técnicas de econometria espacial para uma amostra de apartamentos; Pontes et al. (2011), Paixão (2009) e Rondon & Andrade (2005), que utilizaram o modelo de preços hedônicos para mensurar os custos da criminalidade urbana refletido no preço dos imóveis. Portanto, a construção de índice de preços para imóveis permanece uma lacuna na literatura de modelos hedônicos aplicados a Belo Horizonte.
- d) Por fim, uma vez calculados os índices, podemos fazer uma análise conjuntural do comportamento do preço dos imóveis frente ao ambiente macroeconômico do período.

## 2 Metodologias de cálculo do índice de preços para imóveis

A partir dos trabalhos apresentados no *Workshop on Real Estate Price Indexes*, Diewert (2009) compilou as metodologias que resumem o “estado da arte” na construção de índice de preços para imóveis. As metodologias existentes — método das vendas repetidas (e sua variante, método do valor avaliado), método da estratificação e modelo dos preços hedônicos — diferem no que diz respeito à representatividade da amostra, ao custo de se obter uma base de dados, à forma como lidam com a questão da depreciação e reformas, e à robustez dos resultados.

O método das vendas repetidas (*repeated sale model*), segundo Diewert (2009), consiste em utilizar modelagem econométrica em uma amostra de imóveis que foram vendidos e, posteriormente, revendidos. A vantagem desse método é não necessitar de uma base de dados que contenha informações sobre as diversas características do imóvel, diminuindo os custos de construção da base de dados e tornando o tratamento econométrico menos complexo. A principal desvantagem é utilizar apenas um subconjunto das transações imobiliárias (revendas), podendo esse subconjunto não ser uma amostra representativa do total. Outra desvantagem associada a esse método é a dificuldade de lidar com as questões referentes à depreciação e às reformas. Por fim, esse método é mais adequado para realidades onde as revendas de imóveis sejam frequentes, como é a realidade dos EUA; no caso da Europa e Brasil, por exemplo, onde as transações são menos frequentes pode haver problema de insuficiência da amostra (Gouriéroux & Laferrère 2009, Rozenbaum 2009).

O método do valor avaliado (*sales price appraisal*) é uma variante do método das vendas repetidas, na qual se busca aproveitar informação sobre todas as transações imobiliárias ocorridas em um dado período. Para isso, utiliza o valor avaliado no período subsequente para aqueles imóveis que não foram revendidos. Segundo Diewert (2009), embora esse método contorne o pro-

blema do desperdício de informação inerente ao método das vendas repetidas, a questão de como incorporar a questão da depreciação e dos melhoramentos continua sem solução. Por fim, Diewert (2009) aponta que os resultados desse método são totalmente dependentes da qualidade da avaliação. Estando os valores avaliados distantes dos valores de mercados, os índices calculados terão pouca utilidade.

O método da estratificação (*stratification method*) consiste em estratificar os imóveis em tipos homogêneos quanto às suas características (físicas e locais), e a partir daí, utilizar a média ou mediana para calcular variações temporais nos preços. Para não gerar índices viesados, é necessário que se construam índices específicos para cada característica relevante do imóvel por localização, estando essa metodologia sujeita ao problema de insuficiência da amostra (Diewert 2009). Outro problema inerente a essa metodologia, segundo Diewert (2009), é dificuldade de lidar com as questões referentes à depreciação e melhoramentos.

O método hedônico (*hedonic methods*) consiste em uma regressão em que o preço do imóvel (geralmente em logaritmo) é a variável dependente. As variáveis independentes compreendem as características do imóvel e uma *dummy* para cada período de tempo, sendo o período inicial a categoria básica. O índice de preço é a exponencial do parâmetro estimado para cada *dummy* de tempo. Diewert (2009) cita como principais vantagens desse método o não desperdício de informação, a possibilidade de incorporar a depreciação e os melhoramentos e de se calcular índices específicos por tipo de imóvel e por localização.

Diewert (2009) pondera que existem algumas dificuldades associadas à utilização do modelo de preços hedônicos. Primeiramente, o modelo requer que haja informações sobre os preços e as características físicas e espaciais do imóvel. Não estando essa base disponível nos registros administrativos, construí-la é uma tarefa dispendiosa. Outras dificuldades inerentes às aplicações do modelo de preços hedônicos é o fato de a teoria não especificar a priori a forma funcional da regressão hedônica e nem as características do bem que devem ser incluídas no modelo. Como consequência, diferentes pesquisadores, a partir da mesma amostra de imóveis, podem gerar índices diferentes, na medida em que especificam a forma funcional de forma distinta e/ou incluem um conjunto diferenciado de características no modelo.

Analisando os aspectos favoráveis e desfavoráveis de cada método, Diewert (2009) conclui que o método hedônico é o mais completo. Ao contrário do método das vendas repetidas, o modelo de preços hedônicos utiliza o universo das transações imobiliárias. O método hedônico não depende de informações externas, como o método do valor avaliado, cuja qualidade da informação (exógena) sobre o valor avaliado do imóvel é fundamental para o desempenho da estimativa. Assim como o método da estratificação, o modelo de preços hedônicos é capaz de gerar índices diferenciados por tipo e localização do imóvel. O método hedônico, todavia, é o único que permite incluir, sem apelar para ajustes exógenos, às questões referentes à depreciação e melhoramentos sofridos pelo imóvel ao longo do tempo. Como conclui o autor:

My overall evaluation of the hedonic regression is that it may be probably the best method that could be used in order to construct constant quality price indexes for various types of property, provided that adequate data are available (Diewert 2009, p.105).

### 3 O modelo de preços hedônicos para o mercado imobiliário de Belo Horizonte

#### 3.1 O modelo de preços hedônicos e sua forma funcional

O método utilizado para estimar a variação do preço dos imóveis foi o hedônico com *dummy* de tempo, como proposto em Diewert (2009). A base de dados utilizada para o valor de transação e para as características físicas foi a do ITBI, disponibilizada pela Secretaria de Fazenda da Prefeitura Municipal de Belo Horizonte (SEFAZ/PBH).

O modelo hedônico estabelece uma relação entre o preço e os atributos do imóvel. Para nossos propósitos, a função hedônica pode ser descrita da seguinte forma:

$$P = f(F, I, A, V, C, T) \quad (1)$$

onde  $P$  é o preço do imóvel;  $F$  um conjunto de características físicas do imóvel;  $I$  um conjunto de variáveis que relacionam o preço do imóvel à sua idade;  $A$  um conjunto de características da acessibilidade da localização do imóvel;  $V$  um conjunto de características da vizinhança do imóvel;  $C$  são as características relacionadas à criminalidade na região onde o imóvel se situa; e  $T$  o período de tempo no qual o imóvel foi negociado.

O modelo de preços hedônicos estabelece que o preço marginal (implícito) de cada um dos atributos do bem é calculado a partir da primeira derivada parcial da função hedônica para cada um dos seus componentes. A teoria não estabelece qual a forma funcional e nem quais variáveis devem fazer parte do modelo. Para Griliches (1971), essas questões exigem uma resposta empírica, dependendo dos objetivos do pesquisador e da base de dados que se dispõe.

Utilizaremos neste trabalho um modelo semelhante ao utilizado por Hoesli et al. (1997), no qual a forma da função hedônica é dada por:

$$P = \alpha e^{\left(\sum_{i=1}^{i=n} \beta_i X_i + \sum_{t=1}^{t=j} \delta_t T_t\right)} \quad (2)$$

onde  $X$  é o conjunto das características físicas, de idade, acessibilidade, vizinhança e criminalidade do imóvel; e  $T$  representa o período  $t$  em que a transação ocorreu.  $\beta_i$  é o preço implícito de cada característica  $i$  do imóvel; e  $\delta_i$  é o coeficiente que representa a evolução do preço dos imóveis, em cada período  $t$ , tendo o período inicial como base, após controlado pelas características. A partir de (2), o modelo estimado segue a seguinte especificação:

$$\ln P = \alpha + \sum_{i=1}^{i=n} \beta_i X_i + \sum_{t=1}^{t=j} \delta_t T_t + \epsilon \quad (3)$$

onde  $\ln P$  é o logaritmo natural do preço do bem imóvel;  $\alpha$  o termo da constante;  $\beta_i$  o preço implícito estimado para a  $i$ -ésima característica do imóvel;  $X_i$  cada uma das características física ou ambiental do imóvel;  $\delta_{-t}$  o coeficiente estimado para cada um dos  $t$ 's períodos;  $T$  o período em que se realizou a transação; e  $\epsilon$  o termo do erro associado à regressão. Seguindo Diewert (2009), o índice de preço para cada período  $t$  ( $IP_t$ ), tendo o período inicial como base, é obtido por:

$$IP_t = e^{\delta_t} \quad (4)$$

A variação dos preços entre períodos subsequentes ( $\Delta P_{t,t+1}$ ) esta representada na expressão abaixo:

$$\Delta P_{t,t+1} = e^{\delta_{t+1} - \delta_t} \quad (5)$$

### 3.2 Variáveis incluídas no modelo

A variável dependente: preço da transação

O preço efetivado na transação, em termos nominais, transformado em logaritmo, constante na base de dados do ITBI, foi a variável dependente. Porém, trabalhar com os dados do ITBI requer contornar alguns problemas inerentes a esta base, como a possibilidade de ocorrer subdeclaração ou valores muito distantes da realidade de mercado.

O primeiro cuidado que tem que se tomar é com a possibilidade de haver subdeclaração do preço informado.<sup>1</sup> Compradores e vendedores teriam incentivos para declarar um valor menor do que o efetivamente realizado para pagarem menos impostos (ITBI, no caso dos primeiros, e Imposto de Renda, no caso dos últimos). Para evitar que esse tipo de ação comprometa a arrecadação municipal, as prefeituras, como é o caso da PBH, mantêm um cadastro de preços avaliados para cada imóvel; havendo divergência de valores o ITBI incide sobre o maior dentre o declarado e o avaliado.

Além da subdeclaração, González (1997a) enumera outros problemas que podem ocorrer na informação de preço do ITBI: i) erros de leitura, gravação, transcrição e cálculo dos dados; ii) especificidade do imóvel (alto luxo, localização privilegiada, entre outros) ou da transação (necessidade urgente de venda ou compra); iii) deficiência na coleta, por exemplo, a exclusão da parcela financiada no valor declarado. Para detectar esses problemas, González (1997a) recomenda a utilização de técnicas estatísticas simples, como gráficos de dispersão com valores do preço por m<sup>2</sup> e/ou do logaritmo do preço. Como solução, o autor sugere a remoção dos dados ou a inclusão de *dummy* nos casos em que o pesquisador consegue identificar a razão para o preço discrepante.

A base de dados do ITBI contém, além da informação do preço declarado na transação, o preço avaliado pela prefeitura. Foram retiradas da base de dados todas as transações em que o preço declarado fosse menor que o preço avaliado.<sup>2</sup> Posteriormente, foi feita a análise gráfica de dispersão dos valores do logaritmo e, quando necessária, a análise de dispersão do preço do metro quadrado, como proposto por González (1997a). A Tabela 1 apresenta os limites identificados pela análise gráfica para eliminação de observações com valores discrepantes.

<sup>1</sup>González (1997a) cita outra situação frequente na base do ITBI: a inclusão de valores muito baixos (R\$ 0,01), fato verificado na nossa base. Segundo o autor, isso não é decorrente da subdeclaração, mas devido à correção monetária, dado que o valor constante na guia do ITBI é o nominal de contrato. González (1997a) ainda adverte que mesmo na ausência de subdeclaração o preço declarado no ITBI costuma ser menor que o observado em anúncios de imobiliárias. Isso se deve ao fato de que no preço declarado inclui a margem de negociação — que costuma ocorrer nas transações imobiliárias — sobre o preço ofertado nos anúncios.

<sup>2</sup>No caso do mercado de galpões foram utilizados os filtros constante na Tabela 1, mas não se retirou os dados em que o preço declarado fosse menor que o avaliado, pois isso diminuiu muito o tamanho da amostra, comprometendo a estimação do modelo.

**Tabela 1:** Limites utilizados para eliminação de dados das amostras

Tipo de imóvel	Período	Logartimo	Preço do metro quadrado
Apartamento	1995-2003	$9 < \ln(P) < 14$	Acima de R\$ 4.000,00
Casa	1995-1996	$6 < \ln(P) < 14$	Acima de R\$ 5.000,00
	1997-2003	$7 < \ln(P) < 14$	
Barraco	1995-1996	$7 < \ln(P) < 11$	Acima de R\$ 1.250,00
	1998-2000	$7 < \ln(P) < 11$	Acima de R\$ 1.500,00
	2001-2003	$7 < \ln(P) < 12$	Acima de R\$ 1.500,00
Loja	1995-2003	$8 < \ln(P) < 16$	Abaixo de R\$ 490,00
		Acima de R\$ 4.000,00	
Sala	1995-2003	$8 < \ln(P) < 16$	Abaixo de R\$ 490,00
		Acima de R\$ 4.000,00	
Galpão	1995-2003	$8 < \ln(P) < 15$	Não utilizado
Terreno	1995-2003	$8 < \ln(P) < 16$	Não utilizado

## Variáveis independentes

As variáveis independentes incluídas nos modelos a serem testados estão resumidas na Tabela 2.

Dentre o conjunto de infraestrutura física do imóvel ( $F$ ), temos que espaço é um dos parâmetros que norteiam a decisão dos agentes que atuam no mercado imobiliário, de acordo com os modelos de economia urbana<sup>3</sup> (Abramo 2001, Evans 2003), por isso, espera-se que os coeficientes estimados para área do imóvel e área do terreno tenham sinal positivo. A literatura de preços hedônicos também reconhece o padrão de acabamento como item importante na valorização do preço do imóvel (Hermann & Haddad 2005), ou seja, quanto melhor o padrão de acabamento maior tende a ser o preço do bem.

As variáveis referentes à idade do imóvel ( $I$ ) são idade do imóvel e o quadrado da idade. A idade de um bem durável é uma *proxy* de sua depreciação (Goodman & Thibodeau 1995). Por isso, espera-se uma relação negativa entre preço do bem imóvel e sua idade. O quadrado da idade é incorporado, pois, ao longo dos anos, alguns imóveis passam por reformas, renovações e/ou melhoramentos e, além disso, existem imóveis que são valorizados pelo fato de serem antigos (*vintage effect*) (Goodman & Thibodeau 1995). Esses dois efeitos atuam como um “freio” na depreciação decorrente da idade no valor dos imóveis. Por isso, espera-se um coeficiente positivo para o quadrado da idade do imóvel.

Os atributos de acessibilidade do imóvel ( $A$ ) são representados pela distância ao centro da cidade e concentração de estabelecimentos de comércio e serviços. A acessibilidade é fator-chave na valorização da terra urbana e questão central dos modelos de economia urbana (Abramo 2001, Evans 2003). Segundo esses modelos, existe um gradiente decrescente nos preços da terra urbana, na medida em que esta se afasta do centro da cidade. Belo Horizonte é uma cidade cujo centro tem um importante poder polarizador (Macedo 1998, Amaral et al. 2007, Paixão & Abramo 2008).<sup>4</sup> Por isso, foi inserida a variável

<sup>3</sup>Esses modelos trabalham com a ideia de que ha um *trade off* entre espaço e acessibilidade na escolha de localização, dado que o adquirente de imóvel urbano valoriza ambos.

<sup>4</sup>A Região Central de Belo Horizonte concentra 27,3% dos estabelecimentos de comércio da cidade, 34,2% dos estabelecimentos de serviços e apenas 3,1% da população. O centro de Belo

**Tabela 2:** Conjunto de variáveis independentes que resumem os atributos físicos e espaciais dos imóveis

Conjunto	Variável	Descrição	Unidade de Medida	Tipo	Sinal esperado	Fonte do dado
Infra-estrutura física do imóvel	Imóvel	Área do imóvel	Metro quadrado	Contínua	+	ITBI/SEFAZ/PBH
	Terreno	Área do terreno	Metro quadrado	Contínua	+	ITBI/SEFAZ/PBH
	Luxo	Imóvel padrão luxo		Dummy	+	ITBI/SEFAZ/PBH
	Alto	Imóvel padrão alto	Padrão de	Dummy	+	ITBI/SEFAZ/PBH
	Normal	Imóvel padrão normal	acabamento	Dummy	Categoria básica	ITBI/SEFAZ/PBH
	Baixo	Imóvel padrão baixo	medido pela PBH	Dummy	-	ITBI/SEFAZ/PBH
Idade do imóvel (I)	Popular	Imóvel padrão popular		Dummy	-	ITBI/SEFAZ/PBH
	Idade do imóvel	Número de anos do imóvel desde o ano de lançamento até o ano da transação	Discreta	-	ITBI/SEFAZ/PBH	
	Idadequad	Idade do imóvel elevada ao quadrado	Número de anos ao quadrado	Discreta	+	ITBI/SEFAZ/PBHE
Acessibilidade (A)	Distância	Distância ao centro da cidade Km	Contínua	-	telelistas.net	
	Facilidades	Concentração de estabelecimentos de comércio e serviços	Taxa por 100.000 habitantes por AP	Contínua	+	SEFAZ/PBHE
Características da vizinhança (V)	Hierarquia	Ranking dos bairros em ordem decrescente da renda média familiar	Índice do 0 a 100	Contínua	+	Elaboração própria a partir dos dados do CENSO/IBGE, 2000
	Moradia	Índice de inadequação de moradias	Índice de 0 (para o melhor) e 5 (para o pior)	Discreta	-	Lemos (2003)
	Infra	Índice de inadequação de infraestrutura	Índice de 0 (para o melhor) e 5 (para o pior)	Discreta	-	Lemos (2003)
	Indústria	Concentração de estabelecimentos industriais	Taxa por 100.000 habitantes por AP	Contínua	-	SEFAZ/PBHE
Criminalidade (C)	Homicídios	Concentração da ocorrência de homicídios	Taxa por 100.000 habitantes por AP, para o ano em que o imóvel foi transacionado	Contínua	-	CRISP/FAFICH/UFMG

distância ao centro da cidade. A variável facilidades urbanas foi acrescentada para captar o efeito dos centros locais no preço dos imóveis, uma vez que a proximidade a aglomerações de comércio e serviços, em tese, valoriza o terreno urbano.

O conjunto de variáveis que descrevem a vizinhança do imóvel inclui o indicador de hierarquia socioespacial do bairro, as inadequações de moradia e infraestrutura urbana e um indicador de concentração de estabelecimentos industriais. As características de vizinhança são importantes na formação do valor dos imóveis urbanos. Brueckner et al. (1999) consideram a renda das famílias a melhor *proxy* para a qualidade dos serviços localizados em um bairro. Nesse sentido, foi incluído o indicador de hierarquia socioespacial, e espera-se que quanto maior for o índice para o bairro maior será o preço do imóvel.

As cidades brasileiras são caracterizadas pela grande desigualdade quanto às condições de moradia da população e a oferta de infraestrutura urbana. O índice de inadequação de moradias é uma *proxy* para a presença de vilas e favelas nas proximidades do imóvel. Imóveis localizados próximos a grandes vilas e favelas costumam sofrer um desconto no preço; por isso, espera-se que o indicador tenha sinal negativo.

A ausência de infraestrutura urbana adequada também é fator que desvaloriza o imóvel. Dentre a infraestrutura urbana incluem-se a presença de iluminação pública, saneamento básico e condições adequadas de coleta de lixo, fatores que influenciam diretamente o conforto, a sensação de segurança e a saúde da população local. Por isso, espera-se que o indicador de inadequação de infraestrutura urbana tenha um sinal negativo.

A proximidade de áreas industriais é desvalorizada no mercado imobiliário, uma vez que áreas onde se aglomeram estabelecimentos industriais tendem a ser mais poluídas, distante do centro e com um grande fluxo de pessoas e veículos (gerando engarrafamentos e altos índices de ruídos). Desse modo, espera-se que o sinal do coeficiente estimado para a variável indústria seja negativo.

A variável que representa a criminalidade urbana ( $C$ ) foi a taxa de homicídios por Área de Ponderação ( $AP$ ) por 100.000 habitantes para cada ano analisado. A criminalidade urbana tem sido apontada como fator que contribui para decadência urbana, existindo evidências de que aumentos na taxa de criminalidade são rapidamente capitalizados no mercado imobiliário (Tita et al. 2006). Utilizamos a taxa de homicídios como *proxy* para a criminalidade urbana, uma vez que esse é um crime considerado mais grave pelos participantes do mercado imobiliário.<sup>5</sup> Desse modo, espera-se uma relação negativa entre o coeficiente estimado para homicídios e o preço do imóvel.

As variáveis temporais ( $T$ ), que representam o período em que o imóvel foi negociado, estão representadas na Tabela 3.

A evolução do preço dos bens imóveis será mensurada por ano, uma vez que essa é única informação sobre a data da transação disponível na base de dados utilizada. O índice de preço para cada ano, tendo o ano de 1995 como base, é calculado pelo exponencial do coeficiente estimado, como proposto em

---

Horizonte é ainda destino de 1/3 dos deslocamentos realizados na Região Metropolitana (Paixão & Abramo 2008, p.239-240).

<sup>5</sup>Estudos que utilizaram preços hedônicos para estimar o custo da criminalidade no preço dos imóveis no Brasil encontraram um coeficiente maior para a variável homicídio em relação à variável roubos a transeuntes ou assaltos a mão armada (Rondon & Andrade 2005, Teixeira & Serra 2006, Paixão 2009).

**Tabela 3:** Variáveis que representam o ano de transação do imóvel

Variável	Descrição	Tipo	Fonte do dado
1995	Ano em que o imóvel foi transacionado	Categoria básica	ITBI/SEFAZ/PBH
1996		Dummy	ITBI/SEFAZ/PBH
1997		Dummy	ITBI/SEFAZ/PBH
1998		Dummy	ITBI/SEFAZ/PBH
1999		Dummy	ITBI/SEFAZ/PBH
2000		Dummy	ITBI/SEFAZ/PBH
2001		Dummy	ITBI/SEFAZ/PBH
2002		Dummy	ITBI/SEFAZ/PBH
2003		Dummy	ITBI/SEFAZ/PBH

(5); e a variação anual do preço é estimada a partir do exponencial da diferença entre os coeficientes estimados para os dois anos, como proposto em (6).

## 4 Resultados

As Tabelas (4) e (5) resumem os resultados da regressão, especificada em (3), para cada tipo de imóvel sobre os valores nominais dos imóveis transacionados.<sup>6</sup>

O ajuste do modelo variou de acordo com o tipo de imóvel. O melhor ajuste foi para o mercado de apartamentos, com um  $R^2$ -ajustado de 83,3%. Para os mercados de casa, loja, sala, galpão e terrenos, o ajuste foi satisfatório para dados do tipo *cross-section*, com o  $R^2$ -ajustado em torno dos 50,0%. Para o mercado de barracos o ajuste da regressão foi um pouco pior, com um  $R^2$ -ajustado de 30,9%. A base de dados do ITBI contempla apenas as transações que ocorreram no mercado formal. Grande parte das transações com barracos acontecem no mercado informal, podendo haver problema de insuficiência da amostra.<sup>7</sup>

### 4.1 Variáveis de Controle

O comportamento das variáveis de controle seguiu, na maior parte dos casos, aquele predito no modelo. Dentre as variáveis de Infraestrutura física do imóvel ( $F$ ) área e terreno, seguiram o padrão previsto, com parâmetros estimados positivos e significativos a 5% em todos os mercados — com exceção do mercado de lojas onde o parâmetro para terreno não foi significativo, embora área do terreno seja um item com pouca participação na formação do preço dos diversos tipos de imóveis. Acabamento também seguiu o padrão esperado em todos os mercados, exceto no mercado de casas onde o parâmetro para o acabamento Luxo não foi significativo a 5% e teve um valor estimado menor do que o acabamento Alto. Esse comportamento pode ser reflexo do efeito localização, na qual casas localizadas em áreas centrais ou de expansão imobiliária

<sup>6</sup>No apêndice encontram-se as saídas da regressão para os valores reais, em R\$ de 2003 tendo o IPCA como deflator, dos imóveis transacionados. Nota-se que os parâmetros estimados para as variáveis de controle são idênticos. A constante e os parâmetros estimados para as variáveis temporais, assim como os respectivos desvios padrões estimados, são os que sofrem alterações.

<sup>7</sup>O acesso a imóveis via mercado informal é comum nas cidades brasileiras. Para um entendimento do mercado imobiliário informal, ver Abramo (2003, 2007).

**Tabela 4:** Modelo de Preços Hedônicos para imóveis residenciais: apartamento, casa e barracão

Variáveis independentes	Apartamento		Casa		Barracão	
	Coefficientes	p-value	Coefficientes	p-value	Coefficientes	p-value
Constante	10,189960	0,000000	10,358530	0,000000	9,625652	0,000000
Imóvel	0,005477	0,000000	0,003567	0,000000	0,001908	0,000000
Terreno	-0,000002	0,000000	-0,000018	0,000000	0,000060	0,000000
Luxo	0,231652	0,000100	0,021210	0,837500	-	-
Alto	0,194995	0,000000	0,201048	0,000000	-	-
Baixo	-0,118183	0,000000	-0,208581	0,000000	-0,174486	0,000000
Popular	-0,286931	0,000000	-0,361737	0,000000	-0,363933	0,000000
Idade	-0,016287	0,000000	-0,000950	0,041400	-0,002190	0,456400
Idadequad	0,000101	0,000000	0,000000	0,025300	0,000066	0,216300
Distância	-0,023658	0,000000	-0,014389	0,000000	-0,017208	0,000200
Facilidades	-0,000004	0,000000	0,000082	0,000000	-0,000133	0,000200
Hierarquia	0,007031	0,000000	0,010830	0,000000	0,021217	0,000000
Moradia	-0,000116	0,952500	-0,074759	0,000000	-0,039132	0,309600
Infra	-0,016366	0,000000	0,033151	0,000200	-0,021190	0,489100
Indústria	-0,000203	0,000000	-0,000628	0,005800	0,002174	0,000300
Homicídios	-0,001099	0,000000	-0,001334	0,000200	0,000600	0,592800
1996	0,106957	0,000000	-0,043881	0,202600	0,141379	0,127300
1997	0,214693	0,000000	0,042254	0,152600	0,345191	0,000000
1998	0,317621	0,000000	0,131880	0,000000	0,376158	0,000000
1999	0,255729	0,000000	0,095260	0,001000	0,460115	0,000000
2000	0,279772	0,000000	0,161641	0,000000	0,474599	0,000000
2001	0,293181	0,000000	0,194028	0,000000	0,509911	0,000000
2002	0,350840	0,000000	0,285807	0,000000	0,656009	0,000000
2003	0,401004	0,000000	0,390117	0,000000	0,791087	0,000000
$R^2$ -ajustado	0,8329		0,6180		0,3088	
"F" (Nível de significância)	15885,88 (0,000)		962,3265 (0,000)		39,13996 (0,000)	
Número de observações	73322		13668		1794	

Fonte: Elaboração própria, modelo da eq.(3). Variável dependente:  $\ln(\text{preço})$ ; Desvio padrão de White (consistente com heterocedasticidade).

são valorizadas mais pelo potencial do uso do terreno do que pelo padrão da construção em si.

Dentre as variáveis de Idade ( $I$ ), idade apresentou o comportamento esperado, parâmetro com valor negativo em todos os mercados, embora no mercado de barracos e galpões o valor desse parâmetro foi não significativo a 5%. O comportamento de idade ao quadrado também seguiu o esperado, parâmetro positivo a significativo a 5%, em todos os mercados exceto barracos, onde não foi significativo a 5%, e galpões, onde além de não ter sido significativo a 5% teve o sinal negativo.

Nas variáveis de Acessibilidade ( $A$ ), distância ao centro da cidade teve o comportamento esperado em todos os mercados, exibindo o parâmetro sinal negativo e sendo significativo. Facilidades teve o comportamento esperado — sinal positivo e parâmetro significativo — em todos os mercados, exceto apartamentos, barracos e salas. Nesses três mercados, o parâmetro estimado foi negativo — embora em pequena magnitude — e significativo. Esse resultado pode estar indicando, nesses três mercados, a existência de um *trade off* entre a conveniência de estar localizado em regiões com abundância de comércio e serviços e as desamenidades inerentes a essa proximidade — poluição sonora, grande fluxo de veículos e pessoas, criminalidade etc.

**Tabela 5:** Modelo de Preços Hedônicos para imóveis comerciais (loja, sala e galpão) e terreno

Variáveis independentes	Loja		Sala		Galpão		Terreno	
	Coefficientes	p-value	Coefficientes	p-value	Coefficientes	p-value	Coefficientes	p-value
Constante	10,804990	0,000000	10,123770	0,000000	10,450030	0,000000	10,625020	0,000000
Imóvel	0,001341	0,000000	0,001696	0,000000	0,000245	0,015600	-	-
Terreno	-0,000010	0,371200	0,000057	0,000000	0,000032	0,000700	0,000215	0,000100
Luxo	0,605545	0,000000	0,710364	0,000000	0,424826	0,576200	-	-
Alto	0,360490	0,000000	0,461805	0,000000	0,022477	0,883100	-	-
Baixo	-0,069649	0,064000	0,053356	0,028700	-0,153409	0,015800	-	-
Popular	-0,222136	0,001900	0,009697	0,921500	-0,228019	0,000900	-	-
Idade	-0,011250	0,000700	-0,007738	0,000000	-0,000526	0,912100	-	-
Idadequad	0,000425	0,000000	0,000117	0,019300	-0,000011	0,906900	-	-
Distância	-0,022659	0,003600	-0,060181	0,000000	-0,014733	0,018300	-0,043637	0,000000
Facilidades	0,000017	0,000700	-0,000006	0,001600	0,000032	0,023200	0,000064	0,000300
Hierarquia	0,006530	0,000000	0,004331	0,000000	0,018301	0,000000	0,013952	0,000000
Moradia	-0,085415	0,004200	-0,031210	0,020800	-0,129120	0,008800	-0,079544	0,000000
Infra	0,152007	0,000000	0,143214	0,000000	0,029854	0,485100	0,057481	0,000000
Indústria	-0,001235	0,000000	-0,000035	0,662100	0,000790	0,007800	0,000098	0,763700
Homicídios	0,002621	0,029400	-0,002970	0,000000	0,001844	0,281600	-0,001456	0,011300
1996	0,046318	0,510700	0,066080	0,028000	0,046727	0,643800	0,004320	0,900500
1997	0,180242	0,024800	0,084926	0,008200	0,174674	0,085800	0,050142	0,124900
1998	0,108207	0,109300	0,155606	0,000000	0,201843	0,031900	0,088263	0,003900
1999	0,201305	0,002900	0,186127	0,000000	2,378202	0,000000	0,104491	0,000500
2000	0,198044	0,003400	0,160302	0,000000	0,314296	0,000700	0,096762	0,001500
2001	0,155819	0,035500	0,214978	0,000000	0,386326	0,000100	0,131256	0,000000
2002	0,137855	0,063500	0,196146	0,000000	0,493463	0,000000	0,081936	0,004400
2003	0,170347	0,035100	0,163019	0,000000	1,122728	0,000000	0,189754	0,000000
R <sup>2</sup> -ajustado	0,4538		0,5009		0,4898		0,4548	
“F” (Nível de significância)	108.8425 (0.000000)		279.1363 (0.000000)		118.7487 (0.000000)		559.7346 (0.000000)	
Número de observações	2986		6421		2822		10719	

Fonte: Elaboração própria, modelo da eq.(3).Variável dependente:  $\ln(\text{preço})$ ; Desvio padrão de White (consistente com heterocedasticidade).

No subgrupo das características de Vizinhança (*V*), a variável hierarquia socioespacial apresentou o comportamento esperado. Em todos os mercados o parâmetro estimado foi significativo com sinal positivo. O indicador de moradia apresentou sinal negativo, como esperado, e foi significativo a 5% em todos os mercados, exceto nos mercados de apartamentos e barracos em que o parâmetro não foi significativo a 5%. Ou seja, estar perto de vilas e favelas foi o fator que levou à desvalorização de quase todos os tipos de imóveis, exceto barracos e apartamentos. No caso dos barracos esse comportamento pode ser explicado por estes se localizarem em vilas ou favelas que foram urbanizadas. Já para apartamentos esse resultado surpreendeu o que pode estar associado à localização de grandes complexos de vilas e favelas em áreas ditas “nobres” da cidade.

Inadequação de infraestrutura, por sua vez, só teve o parâmetro estimado significativo e negativo — como esperado — no mercado de apartamentos. Para barracões, o parâmetro estimado foi negativo e não significativo. Nos mercados de casas, terrenos, salas e lojas, surpreendentemente, o parâmetro estimado, além de positivo, foi significativo. No caso dos terrenos o sinal positivo pode ser reflexo da expectativa de instalação de infraestrutura no futuro. No caso das salas e lojas os coeficientes estimados foram altos (0,15 e 0,14 respectivamente). Para as casas o valor do parâmetro estimado foi menor (0,03). Esses resultados podem estar refletindo alguma peculiaridade do mercado de imóveis comerciais — salas e lojas — e do mercado de casas não captadas pelas hipóteses aqui adotadas.

O parâmetro estimado para a variável indústria seguiu o esperado, sinal negativo, na maior parte dos mercados, exceto nos mercados de barracos e galpões. Porém, para casas o parâmetro estimado foi significativo somente a 10%, já para o mercado de salas o parâmetro estimado não foi significativo nem a 10%. Para barracos e galpões, os parâmetros estimados foram positivos e significativos. No caso dos barracos, o sinal positivo pode estar refletindo a escolha locacional dos operários da indústria que pagam um prêmio para residirem próximos aos seus locais de trabalho. Desse modo, no mercado de barracos, estar próximo às indústrias é, para certos tipos de família, tido mais como uma economia de tempo e de custo de transporte do que uma desamenidade. No mercado de galpões, o prêmio pago para se localizar próximo a indústria pode ser reflexo, tanto da economia de custo de transporte entre o local de fabricação e armazenamento dos produtos quanto do fato das aglomerações industriais estarem perto de grandes rodovias, facilitando o fluxo dos produtos das grandes redes de atacadistas.

Por fim, homicídios, variável que representa a Criminalidade urbana (*C*), teve o comportamento esperado, parâmetro estimado com sinal negativo e significativo a 5%, em todos os mercados, exceto barracos, galpões e lojas. No mercado de barracos e galpões, o parâmetro estimado foi positivo, mas não significativo. Para o mercado de lojas, o parâmetro estimado, além de positivo, foi significativo. Esse resultado pode ser consequência do fato de que áreas com grandes concentrações de comércio são também áreas onde costumam ocorrer grande número de homicídios. Essas áreas costumam ser vazias à noite, dado que a maior parte dos estabelecimentos comerciais encontra-se fechada e o fluxo de pessoas se torna pouco intenso, o que torna essas áreas mais suscetíveis a esse tipo de crime.

## 4.2 Índice de preços para imóveis

Os coeficientes estimados para as *dummies* de tempo permitem o cálculo dos índices de preços ( $IP_t$ ) e das variações anuais nos preços ( $\Delta P_{t,t+1}$ ). Para apartamentos e salas, todos os coeficientes estimados foram significativos. No caso dos barracos, o primeiro ano não foi significativo a 5%. Nos mercados de casas, terrenos e galpões, os parâmetros estimados para os dois primeiros anos não foram significativos a 5%, embora no caso de galpões o segundo ano tenha sido significativo a 10%. Para as lojas, os parâmetros estimados para o primeiro e o terceiro ano não foram significativos. Como os coeficientes estimados para essa variável representam a variação de preço a partir do ano base, é esperado que com o passar dos anos eles se tornem cada vez mais significativos (Hoesli et al. 1997, p.102). A Tabela 6 apresenta a variação real anual do preço dos imóveis (equação 5), calculada a partir dos coeficientes estimados (equação 3) deflacionados pelo Índice de Preços ao Consumidor Amplo (IPCA).<sup>8</sup> A Tabela 7 apresenta algumas variáveis macroeconômicas relevantes para o entendimento do mercado imobiliário.

**Tabela 6:** Taxa de variação anual real do preço dos imóveis em Belo Horizonte 1995-2003

Período	Apartamento	Casa	Barraco	Loja	Sala	Galpão	Terreno
1996	1,58	-12,64	5,14	-4,40	-2,49	-4,36	-8,33
1997	5,85	3,59	16,52	8,66	-3,16	8,01	-0,50
1998	9,03	7,59	1,46	-8,47	5,57	1,08	2,19
1999	-13,72	-11,51	-0,17	0,75	-5,36	709,08	-6,70
2000	-3,34	0,84	-4,26	-5,95	-8,04	-88,02	-6,36
2001	-5,87	-4,08	-3,79	-10,96	-1,90	-0,19	-3,86
2002	-5,86	-2,59	2,84	-12,72	-12,79	-1,09	-15,41
2003	-3,80	1,55	4,72	-5,49	-11,49	71,66	1,91
Média* 1996-1998	5,44	-0,88	7,52	-1,67	-0,10	1,45	-2,32
Média* 1999-2003	-6,59	-3,27	-0,19	-6,99	-8,00	10,44	-6,26

Fonte: Elaboração própria. Taxa de variação calculada pela eq.(5) e deflacionadas pelo IPCA; \* Média geométrica.

O período entre 1996 e 2003 foi marcado pela consolidação da estabilidade de preços. Em janeiro de 1999, houve a maxidesvalorização do Real, que dividiu o período em dois. Entre 1996 e 1998, a estabilidade de preços foi mantida por meio da âncora cambial, um regime de câmbio semifixo, que, conjugada com as boas condições do mercado financeiro internacional, permitiu com que o Real se apreciasse. Os déficits em Transações Correntes, oriundos da taxa de câmbio apreciada, eram cobertos por capitais externos atraídos pela alta taxa de juros doméstica (Selic) e pela grande liquidez do mercado internacional. Nesse contexto, a despeito da taxa juros alta, o mercado de crédito bancário e o mercado de capitais expandiram-se (Hermann 2010).

O comportamento do mercado imobiliário, nessa primeira fase, foi diverso. Apartamentos, casas e barracos apresentaram aumento real de preços. O efeito negativo da alta taxa de juros — média de 20,44% no período — foi mais que contrabalanceado pelo crescimento da renda,<sup>9</sup> 3,98% em média, e

<sup>8</sup>No apêndice encontra-se a valorização dos imóveis calculada a partir da regressão com os valores das transações em R\$ de 2003 tendo o IPCA como deflator. Nota-se que o valor estimado é basicamente o mesmo, com algumas diferenças a partir da terceira casa decimal.

<sup>9</sup>O PIB, entre 1996-1998, cresceu em ritmo menor que no período 1999-2003. Porém, esten-

**Tabela 7:** Indicadores macroeconômicos — Brasil — 1996-2003

Período	PIB(1)	Inflação (IPCA)	SELIC(2)	Renda(3)	Volume de Crédito imobiliário(4)	Variação da taxa de câmbio(5)
1996	2,15	9,56	27,41	10,65	-1,79	3,35
1997	3,38	5,22	18,60	2,01	-2,75	3,67
1998	0,04	1,66	26,69	-0,45	3,96	7,60
1999	0,25	8,94	15,28	-5,5	-4,41	47,39
2000	4,31	5,97	10,81	-1,19	0,20	-1,55
2001	1,31	7,67	8,96	-3,37	-41,88	21,18
2002	2,66	12,53	5,90	-4,02	-40,53	13,18
2003	1,15	9,3	12,85	-11,88	-6,24	-1,62
Média 1996-1998 (6)	1,84	5,43	20,44	3,98	-0,24	4,86
Média 1999-2003 (6)	1,93	8,86	10,72	-5,26	-20,86	14,38

Fonte: elaboração própria, dados Ipeadata.

<sup>1</sup> Taxa de crescimento real (R\$ de 1980).

<sup>2</sup> Taxa mensal Over/Selic anualizada e corrigida pelo Índice de Preços ao Consumidor Amplo (IPCA).

<sup>3</sup> Taxa de crescimento da renda real habitual das pessoas ocupadas das Regiões Metropolitanas. Até o ano de 2002 foi utilizada a série antiga da Pesquisa Mensal do Emprego (PME). A partir de 2002 foi utilizada a série nova.

<sup>4</sup> Taxa de crescimento dos saldos das operações com pessoas físicas e cooperativas habitacionais deflacionada pelo IPCA.

<sup>5</sup> Variação da taxa de câmbio (R\$/US\$) deflacionada pelo IPCA e pelo (Consumer Price Index) CPI dos EUA.

<sup>6</sup> Média geométrica.

pela relativa estabilidade no volume de crédito habitacional (queda média de 0,24% no período). Barraco foi o imóvel com maior alta no período, 7,52%, em média, o que pode ser atribuído pelo ganho de renda real das camadas mais pobres da população propiciado pela diminuição da inflação. No caso de apartamentos, a valorização foi mais modesta, 5,44%, e no caso das casas uma pequena desvalorização real foi registrada, 0,88%, o que pode ser atribuído ao comportamento do crédito imobiliário, que teve uma pequena queda no período.

Os preços das lojas, salas e terrenos sofreram pequena queda em termos reais, entre 1995-1998. No caso de lojas e salas, embora a renda da população estivesse registrando aumento, o alto patamar da taxa de juros pode ter feito com que lojistas e prestadores de serviços preferissem alugar estabelecimentos em vez de comprá-los. No mercado de terrenos, as taxas de juros elevadas, combinada com a pequena queda no volume de crédito habitacional, inibiram o investimento na construção de novas unidades residenciais, por parte das construtoras. Como o terreno urbano é insumo básico para a produção de novas moradias, a pouca disposição dos empresários em investirem na construção civil de moradias fez com que o valor real dos terrenos caísse (2,32% em média).

A partir de 1997, as condições do mercado financeiro internacional começaram a se deteriorar com ataques especulativos contra moedas, que resultaram nas crises cambiais, nos países do Leste da Ásia (1997-1998), Rússia (1998) e Brasil (1999). Em 1998, a economia brasileira já exibiu os efeitos da crise, com um crescimento do PIB de 0,04%. A âncora cambial foi mantida, a despeito de o Real passar por ataques especulativos durante 1998, o que resultou em grandes perdas de reservas para o Brasil. Em janeiro de 1999, frente à grande perda de reservas propiciada por um ataque especulativo, a âncora cambial foi abandonada de forma traumática — não por uma decisão planejada de política econômica, mas por conta da conjuntura econômica internacional (Modenesi 2005). A partir de 1999, o Brasil passou para um regime de câmbio flutuante, e a âncora nominal passou a ser a meta de inflação.

Os mercados de apartamentos e de casas sofreram diretamente o impacto da maxidesvalorização do Real, com desvalorização real de 13,72% e 11,51%, em 1999, respectivamente. A mudança no regime cambial gerou incertezas nos agentes econômicos quanto ao comportamento futuro da inflação. O PIB, que já havia crescido pouco, em 1998, voltou a ter um crescimento muito pequeno, 0,25%, em 1999. O crédito imobiliário encolheu em 4,41% em valores reais, em 1999, frente a um crescimento real de 3,96%, em 1998. A conjunção desses fatores — incerteza, baixo crescimento e queda no volume de crédito habitacional — fizeram diminuir a demanda por imóveis residenciais em 1999 e, conseqüentemente, quem precisou vender imóveis, nessa conjuntura, teve que aceitar grandes descontos.

O período entre 1999-2003 foi marcado pelas instabilidades no mercado financeiro internacional e pelo baixo crescimento da economia mundial, o que acarretou em um baixo crescimento da economia brasileira, com um refluxo no mercado de crédito e de capitais (Hermann 2010). A política econômica

---

dendo a análise para o período 1995-1998, a taxa de crescimento média do PIB real entre 1995-1998 fica em 3.32%, evidenciando que o período 1995-1998 foi marcado por maior crescimento que o período 1999-2003. O ano de 1995 não foi incluído na análise por não estarem disponíveis os dados de transações imobiliárias de 1994, tornando impossível a estimação da valorização imobiliária entre 1994 e 1995.

passou a ser desenhada seguindo o “tripé” macroeconômico: metas de inflação, metas de superávit primário e regime de câmbio flutuante.

Os imóveis residenciais — apartamentos, casas, barracos — sofreram desvalorização real no período,  $-6,60\%$ ,  $-3,27\%$  e  $-0,19\%$ , em média, respectivamente. Entre os fatores que explicam esse comportamento esta a queda na renda real média dos trabalhadores,  $5,26\%$ , a taxa real básica de juros elevada, embora menor que no período anterior,  $10,72\%$ , e a forte queda no montante real de crédito habitacional,  $20,86\%$ . Renda baixa, juros elevados e queda no crédito imobiliário foram fatores que retraíram a demanda por imóveis, levando à desvalorização real observada. A desvalorização foi menor no mercado de barracos, provavelmente pelas aquisições nesse mercado serem menos dependentes do financiamento habitacional, sendo os preços menos sensíveis a variações no montante de crédito habitacional e na taxa de juros.

Lojas e salas registraram desvalorização ainda maior em comparação ao período anterior,  $6,99\%$  e  $8,00\%$ , respectivamente. A combinação de baixo crescimento econômico, queda na renda real dos trabalhadores e taxa de juros elevada desestimularam comerciantes e prestadores de serviços a expandirem suas atividades e, em consequência, adquirirem novas unidades. Desse modo, a demanda por esses tipos de imóveis retraiu-se ainda mais, o que explica a maior desvalorização. No caso dos terrenos, a desvalorização média, entre 1999 e 2003, foi de  $6,26\%$ . Esse resultado reflete a baixa expectativa dos empresários da construção civil quanto à viabilidade dos empreendimentos em um cenário de baixo crescimento econômico, queda na renda dos trabalhadores e, principalmente, alta taxa de juros e pouco crédito disponível para habitação.

Por fim, os índices de preços foram muito erráticos para o mercado de galpões. Entre 1996-1998, estes apresentaram valorização real de  $1,45\%$ . No período 1999-2003, essa média passou para  $10,44\%$ . Chama atenção nesse mercado, anos de valorização muito expressiva como 1999,  $709,8\%$ , e 2003,  $71,66\%$ , assim como a grande desvalorização de  $88,02\%$ , em 2000, posterior à grande valorização de 1999. Isso pode indicar problemas de transcrição de dados na base, não detectados pelos filtros propostos, ou alguma peculiaridade do mercado desse tipo de imóvel — as variáveis escolhidas podem ter sido insuficientes e/ou inadequadas para a análise, ou mesmo o modelo de preços hedônicos pode não ser a ferramenta mais adequada para se calcular índice de preços para esse mercado específico.

## 5 Conclusão

Elaborar índice de preços que deem conta da valorização imobiliária é uma tarefa importante, tanto para os agentes que atuam no mercado imobiliário, como para participantes do mercado financeiro, autoridades monetárias e *policy makers*. O imóvel é um bem heterogêneo quanto aos atributos que o compõe e, por isso, qualquer cálculo de variação do preço tem que levar em conta essa heterogeneidade. A literatura apresenta quatro métodos que mensuram índice de preços de imóveis, levando em conta sua heterogeneidade: o método de vendas repetidas, o método do valor avaliado, os métodos estratificados e o método hedônico.

O método hedônico é recomendado por aproveitar informações sobre o universo dos imóveis transacionados em cada período e por incorporar, sem

ajustes exógenos, a questão da depreciação, reformas, melhoramentos e renovações. No caso brasileiro, as características do mercado imobiliário: poucas transações e ausência de uma base de dados sobre hipotecas torna difícil a aplicação do método das vendas repetidas. Por isso, foi adotado na análise o método dos preços hedônicos.

A base de dados utilizada foi a das transações imobiliárias ocorridas entre 1995 e 2003 para bens imóveis. Foram incluídos na análise os apartamentos, casas, barracos, lojas, salas, galpões e terrenos. As variáveis explicativas foram selecionadas a partir dos conjuntos de características físicas e espaciais que influenciam no valor do imóvel. A escolha foi realizada tendo como base as teorias de economia urbana, a literatura sobre preços hedônicos e a realidade da estruturação urbana brasileira.

O modelo teve melhor adequação ao mercado de apartamentos, embora o ajuste para os mercados de casas, lojas, salas, galpões e terrenos tenham sido satisfatórios. Para o mercado de barracos, o ajuste foi menor, provavelmente devido à parcela significativa das transações com esse tipo de imóvel se dar no mercado informal. Área do imóvel, padrão de acabamento, distância ao centro e hierarquia socioespacial foram fatores que influenciaram a formação do preço do bem nos diversos mercados analisados. O mercado de barracos apresentou algumas peculiaridades, como ter o coeficiente estimado para concentração de estabelecimentos industriais positivo, resultado que pode refletir o prêmio pago pelos empregados da indústria para residirem próximo ao local trabalho. No mercado de casas, imóveis com padrão de acabamento luxo foram menos valorizados que imóveis de padrão alto. Esse resultado pode ser o reflexo da valorização dos terrenos de casas localizadas próximas ao centro ou em áreas de expansão imobiliária, que podem ser aproveitados para a construção de edifícios de apartamentos ou comerciais.

Os coeficientes estimados para as *dummy* de tempo foram, em geral, significativos, e a partir destes, foi possível calcular os índices de preços e as variações de preço dos bens imóveis. Os índices de preços estimados para o mercado de galpões foram muito erráticos, levantando dúvidas sobre a adequação do método hedônico para gerar índices de preços para esse tipo de imóvel.

No período anterior à maxidesvalorização cambial (1996-1998), apartamentos, casas e barracos valorizaram em termos reais. Em 1999, ano da mudança no regime cambial, foi registrada uma grande perda de valor real para apartamentos e casas decorrentes do cenário de incerteza. No caso dos barracos, a perda de valor foi menor. Entre 1999 e 2003, apartamentos, casas e barracos perderam valor em termos reais. A política de juros altos, a diminuição no rendimento da população, o baixo crescimento econômico e a queda no montante de crédito habitacional são os fatores que explicam a desvalorização real dos imóveis residenciais no período.

Lojas e salas desvalorizaram, em termos reais, no período 1996-2003, devido principalmente ao fraco desempenho da economia, o que inibiu os empresários de investirem em novas unidades de negócios. O comportamento dos terrenos foi semelhante do verificado para as salas e lojas. Como o terreno é o insumo básico para a construção de novas unidades residenciais ou empresariais, o comportamento do preço dos terrenos é reflexo da falta de perspectiva dos empresários da construção civil.

Pesquisas futuras que aumentem o escopo temporal seriam importantes para entender melhor as tendências e os ciclos do mercado de imóveis belo-

horizontino (ou brasileiro, na medida em que forem estendidos para outras localidades) e as relações entre preço dos imóveis, o ambiente macroeconômico e o rendimento de outros ativos reais e financeiros. A inclusão do mercado imobiliário informal seria importante para o melhor entendimento da variação dos preços dos barracos. Por fim, trabalhos como este podem ser estendidos para outras cidades brasileiras, servindo como base para no futuro se construir um índice de preços para bens imóveis no Brasil.

### Referências Bibliográficas

Abramo, P. (2001), *Mercado e ordem urbana: do caos à teoria da localização residencial*, Rio de Janeiro.

Abramo, P. (2003), Teoria econômica da favela. quatro notas sobre a localização residencial dos pobres e o mercado imobiliário informal, *in* 'A cidade da informalidade. O desafio das cidades latino-americanas', Sette Letras/Faperj, Rio de Janeiro.

Abramo, P. (2007), 'A cidade com-fusa: a mão inoxidável do mercado e a produção da estrutura urbana nas grandes metrópoles latino-americanas', *Revista Brasileira de Estudos Urbanos e Regionais* 9(2), 25–54.

Aguirre, A. & Faria, D. v. (1997), 'A utilização de preços hedônicos na avaliação social de projetos', *Revista Brasileira de Economia* 51(3), 391–411.

Aguirre, A. & Macedo, P. (1996), Estimativa de preços hedônicos para o mercado imobiliário de belo horizonte, *in* ANPEC, ed., '8 Encontro Nacional de Econometria', Vol. 1, Aguas de Lindoia, pp. 1–16.

Albuquerque, E., Melo, A. & Souza, H. (2007), Ativo ambiental e preço de imóvel em Recife: um estudo exploratório a partir da utilização do método dos preços hedônicos, *in* 'VII Encontro da Sociedade Brasileira de Economia Ecológica', Fortaleza.

Amaral, P., Simões, R. & Hermeto, A. (2007), A centralidade é o centro: uma análise do setor de serviços em bh, *in* 'V Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos', Enaber, Recife.

Amrein, C. (2010), Capital humano e capital urbano: o impacto das escolas nos preços dos imóveis no município de são paulo, Dissertação (mestrado em ciências econômicas), Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo.

Batalhone, S., Nogueira, J. & Mueller, B. (2002), *Economics of air pollution: hedonic price model and smell consequences of sewage treatment plants in urban areas*, Brasília, DF: Universidade de Brasília.

Brueckner, J., Thisse, J. & Zenou, Y. (1999), 'Why is central paris rich and downtown detroit poor? an amenity-based theory', *European Economic Review* 43, 91–107.

Clark, D., s. . (2006), 'Externality effects on residential property values: the example of noise disamenities', *Growth and Change* 37(3), 460–488.

Diewert, W. E. (2009), The paris oecd-imf workshop on real estate price indexes: conclusions and future directions, in W. E. E. Diewert, ed., 'Price and Productivity Measurement: volume 1 — housing', Trafford Press.

Evans, A. (2003), 'The development of urban economics in the twentieth century', *Regional Studies, Brighton* 37, 521–9.

Faria, R. e. a. (2008), 'Uma aplicação do método de preços hedônicos no setor de saneamento: o projeto de são bento do sul-sc', *Planejamento e Políticas Públicas* (31), 116–127.

Freeman, A. (1979), 'Hedonic prices, property values and measuring environmental benefits: a survey of the issues', *The Scandinavian Journal of Economics* 81(2), 154–173.

Furtado, B. (2007), 'Mercado imobiliário e a importância das características locais: uma análise quantílico-espacial de preços hedônicos em belo horizonte', *Análise Econômica* 47, 71–98.

Gatzlaff, D. H. & Ling, D. C. (1994), 'Measuring changes in local house prices: an empirical investigation of alternative methodologies', *Journal of Urban Economics* 35, 221–244.

Gibbons, S. & Machin, S. (2003), 'Valuing english primary schools.', *Journal of Urban Economics* 53(2), 197–219.

González, M. A. S. (1997a), Fonte alternativa de informação para estudos infra-urbanos, in 'ITBI. Encontro Nacional da ANPUR, 6', Recife, pp. 129–147.

González, M. A. S. (1997b), 'Variação qualitativa e índice de preços na análise do comportamento recente dos aluguéis residenciais em porto alegre (1994-1997)', *Análise Econômica* 28, 133–147.

Goodman, A. C. & Thibodeau, T. G. (1995), 'Age-related heteroskedasticity in hedonic house price equations', *Journal of Housing Research* 6(1), 25–42.

Gouriéroux, C. & Laferrère, A. (2009), 'Managing hedonic housing price indexes: the french experience', *Journal of Housing Economics* 18.

Griliches, Z. (1971), Introduction: hedonic price indexes revisited, in Z. E. Griliches, ed., 'Prices indexes and quality change: studies in new methods of measurement', Cambridge: Harvard University Press.

Hermann, B. M. & Haddad, E. A. (2005), 'Mercado imobiliário e amenidades urbanas: a view through the window', *Estudos Econômicos* 35(2), abr./jun.

Hermann, J. (2010), 'Liberalização e desenvolvimento financeiro: lições da experiência brasileira no período 1990-2006', *Economia e Sociedade* 19(2).

Hoesli, M., Giaccotto, C. & Favarger, P. (1997), 'Three new real estate price indexes for geneva, switzerland', *Journal of Real Estate Finance and Economics* 15(1), 93–109.

Lemos, C. B. (2003), Questões urbanísticas e socioculturais do município de belo horizonte-mg, Lemos, c. b. (coord.), CEDEPLAR/UFMG, Belo Horizonte.

**URL:** Disponível em: <<http://www.cedeplar.ufmg.br/pesquisas/pbh/arquivos/mod2parte1.pdf>>. Acesso em: mar. 2010.

Li, W., Prud'Homme, M. & Yu, K. (2006), Studies in hedonic resale housing price indexes., Real Estate Price Indexes Paper 10, OECD-IMF Workshop, Paris.

**URL:** Disponível em: <<http://economics.ca/2006/papers/0406.pdf>>. Acesso em: ago. 2009.

Macedo, P. B. (1998), 'Hedonic price with spatial effects: an application to the housing market of belo horizonte, brazil', *Revista Brasileira de Economia* 52(1), 63–81.

Maurer, R., Pitzer, M. & Sebastian, S. (2004), 'Hedonic price indices for the paris housing market', *Allgemeines Statistisches Archiv* 88, 303–326.

Miranda, R. & Amaral, H. (2009), Precificação de imóveis residenciais: uma abordagem hedônica-multinível, in 'Anais do IX Encontro Brasileiro de Finanças', São Leopoldo.

Modenesi, A. (2005), *Regimes monetários: teoria e experiência do Real*, Barueri: Manole.

Paixão, L. (2009), 'O impacto da violência no preço dos imóveis comerciais de belo horizonte: uma abordagem econômica', *Revista de Economia Aplicada* 13(1), 125–152.

Paixão, L. (2010), 'Externalidades de vizinhança, estruturação do espaço intraurbano e preço dos imóveis: evidências para o mercado de apartamentos de belo horizonte', *Ensaio FEE* 31(1), 235–258.

Paixão, L. & Abramo, P. (2008), 'Os vetores de expansão da atividade imobiliária em belo horizonte: 1994-2003', *Nova Economia, Belo Horizonte* 18(2), 229–263.

Pontes, E., Paixão, L. & Abramo, P. (2011), 'O mercado imobiliário como revelador das preferências espaciais: uma análise do impacto da criminalidade urbana no preço de apartamentos em belo horizonte', *Revista de Economia Contemporânea* 15(1), 171–197.

Rondon, V. & Andrade, M. (2005), 'Uma estimação dos custos da criminalidade em belo horizonte', *Ensaio FEE* 26(2), 829–854.

Rozenbaum, S. (2009), IMPA — Índice Municipal de Preços de Apartamentos: propostas de metodologia, Doutorado em administração de empresas, Programa de Pós-Graduação em Administração de Empresas, PUC-RIO, Rio de Janeiro.

Rozenbaum, S. & Macedo-Soares, T. D. L. V. A. (2007), 'Proposta para construção de um índice local de preços de imóveis a partir dos lançamentos imobiliários de condomínios residenciais', *Revista de Administração Pública* 41(6), 1069–1094.

Sheppard, S. (2010), Measuring the impact of culture using hedonic analysis, Technical report, Williamstown: Center for Creative Community Development.

URL: Disponível em: < <http://web.williams.edu/Economics/ArtsEcon/library/pdfs/UsingHedonicAnalysis2010.pdf>>. Acesso: jun. 2013.

Teixeira, E. & Serra, M. (2006), 'O impacto da criminalidade no valor de locação de imóveis: o caso de Curitiba', *Economia e Sociedade* 15(1), 175–207.

Tita, G. E., Petras, T. L. & Greenbaum, R. T. (2006), 'Crime and residential choice: a neighborhood level analysis of the impact of crime on housing prices', *Journal of Quantitative Criminology* 22, 299–317.

## Apêndice A

**Tabela A.1:** Modelo de Preços Hedônicos para imóveis residenciais: apartamento, casa e barracão

Variáveis independentes	Apartamento		Casa		Barracão	
	Coefficientes	p-value	Coefficientes	p-value	Coefficientes	p-value
Constante	10,77322	0,000000	10,941780	0,000000	10,208910	0,000000
Imóvel	0,005477	0,000000	0,003567	0,000000	0,001908	0,000000
Terreno	-0,000002	0,000000	-0,000018	0,000000	0,000060	0,000000
Luxo	0,231652	0,000100	0,021210	0,837500	-	-
Alto	0,194995	0,000000	0,201048	0,000000	-	-
Baixo	-0,118183	0,000000	-0,208581	0,000000	-0,174486	0,000000
Popular	-0,286931	0,000000	-0,361737	0,000000	-0,363933	0,000000
Idade	-0,016287	0,000000	-0,000950	0,041400	-0,002190	0,456400
Idadequad	0,000101	0,000000	0,000000	0,025300	0,000066	0,216300
Distância	-0,023658	0,000000	-0,014389	0,000000	-0,017208	0,000200
Facilidades	-0,000004	0,000000	0,000082	0,000000	-0,000133	0,000200
Hierarquia	0,007031	0,000000	0,010830	0,000000	0,021217	0,000000
Moradia	-0,000116	0,952500	-0,074759	0,000000	-0,039132	0,309600
Infra	-0,016366	0,000000	0,033151	0,000200	-0,021190	0,489100
Indústria	-0,000203	0,000000	-0,000628	0,005800	0,002174	0,000300
Homicídios	-0,001099	0,000000	-0,001334	0,000200	0,000600	0,592800
1996	0,015609	0,005500	-0,135228	0,000100	0,050031	0,589400
1997	0,072421	0,000000	-0,100018	0,000700	0,202919	0,016000
1998	0,158935	0,000000	-0,026806	0,364700	0,217472	0,015400
1999	0,011418	0,026200	-0,149051	0,000000	0,215804	0,010900
2000	-0,022569	0,000000	-0,140700	0,000000	0,172259	0,040900
2001	-0,083092	0,000000	-0,182245	0,000000	0,133638	0,129700
2002	-0,143485	0,000000	-0,208518	0,000000	0,161684	0,059500
2003	-0,182252	0,000000	-0,193139	0,000000	0,207831	0,019000
R <sup>2</sup> -ajustado	0,833285		0,621018		0,222743	
"F" (Nível de significância)	15934,85 (0,0000)		974,714 (0,0000)		25,46811 (0,0000)	
Número de observações	73322		13668		1794	

Fonte: Elaboração própria, modelo da eq.(3). Variável dependente:  $\ln(\text{preço})$  em R\$ de 2003; Desvio padrão de White (consistente com heterocedasticidade).

**Tabela A.2:** Modelo de Preços Hedônicos para imóveis comerciais (loja, sala e galpão) e terreno

Variáveis independentes	Loja		Sala		Galpão		Terreno	
	Coefficientes	p-value	Coefficientes	p-value	Coefficientes	p-value	Coefficientes	p-value
Constante	11,388250	0,000000	10,707030	0,000000	11,033290	0,000000	11,208280	0,000000
Imóvel	0,001341	0,000000	0,001696	0,000000	0,000245	0,015600	-	-
Terreno	-0,000010	0,371200	0,000057	0,000000	0,000032	0,000700	0,000215	0,000100
Luxo	0,605545	0,000000	0,710364	0,000000	0,424826	0,576200	-	-
Alto	0,360490	0,000000	0,461805	0,000000	0,022477	0,883100	-	-
Baixo	-0,069649	0,064000	0,053356	0,028700	-0,153409	0,015800	-	-
Popular	-0,222136	0,001900	0,009697	0,921500	-0,228019	0,000900	-	-
Idade	-0,011250	0,000700	-0,007738	0,000000	-0,000526	0,912100	-	-
Idadequad	0,000425	0,000000	0,000117	0,019300	-0,000011	0,906900	-	-
Distância	-0,022659	0,003600	-0,060181	0,000000	-0,014733	0,018300	-0,043637	0,000000
Facilidades	0,000017	0,000700	-0,000006	0,001600	0,000032	0,023200	0,000064	0,000300
Hierarquia	0,006530	0,000000	0,004331	0,000000	0,018301	0,000000	0,013952	0,000000
Moradia	-0,085415	0,004200	-0,031210	0,020800	-0,129120	0,008800	-0,079544	0,000000
Infra	0,152007	0,000000	0,143214	0,000000	0,029854	0,485100	0,057481	0,000000
Indústria	-0,001235	0,000000	-0,000035	0,662100	0,000790	0,007800	0,000098	0,763700
Homicídios	0,002621	0,029400	-0,002970	0,000000	0,001844	0,281600	-0,001456	0,011300
1996	-0,045030	0,522500	-0,025268	0,400800	-0,044620	0,658900	-0,087028	0,011800
1997	0,037970	0,636200	-0,057346	0,074200	0,032403	0,749900	-0,092129	0,004800
1998	-0,050479	0,455000	-0,003080	0,919000	0,043157	0,646300	-0,070423	0,021400
1999	-0,043006	0,524800	-0,058184	0,047900	2,133890	0,000000	-0,139820	0,000000
2000	-0,104296	0,122600	-0,142039	0,000000	0,011955	0,897900	-0,205578	0,000000
2001	-0,220454	0,002900	-0,161295	0,000000	0,010053	0,917300	-0,245017	0,000000
2002	-0,356470	0,000000	-0,298179	0,000000	-0,000862	0,992800	-0,412389	0,000000
2003	-0,412909	0,000000	-0,420237	0,000000	0,539472	0,000000	-0,393502	0,000000
R <sup>2</sup> -ajustado	0,44961		0,509937		0,477973		0,496583	
"F" (Nível de significância)	107.0184 (0.0000)		291.4504 (0.0000)		113.3015 (0.0000)		661.7804 (0.0000)	
Número de observações	2986		6421		2822		10719	

Fonte: Elaboração própria, modelo da eq.(3).Variável dependente:  $\ln(\text{preço})$  em R\$ de 2003; Desvio padrão de White (consistente com heterocedasticidade).

**Tabela A.3:** Taxa de variação anual real do preço dos imóveis em Belo Horizonte 1996-2003

Período	Apartamento	Casa	Barraco	Loja	Sala	Galpão	Terreno
1996	1,57	-12,65	5,13	-4,40	-2,50	-4,36	-8,33
1997	5,85	3,58	16,52	8,65	-3,16	8,01	-0,51
1998	9,04	7,60	1,47	-8,47	5,58	1,08	2,19
1999	-13,72	-11,51	-0,17	0,75	-5,36	709,08	-6,70
2000	-3,34	0,84	-4,26	-5,94	-8,04	-88,02	-6,36
2001	-5,87	-4,07	-3,79	-10,97	-1,91	-0,19	-3,87
2002	-5,86	-2,59	2,84	-12,72	-12,79	-1,09	-15,41
2003	-3,80	1,55	4,72	-5,49	-11,49	71,66	1,91
Média* 1996-1998	5,44	-0,89	7,52	-1,67	-0,10	1,45	-2,32
Média* 1999-2003	-6,60	-3,27	-0,19	-6,99	-3,27	10,44	-6,26

Fonte: Elaboração própria. Taxa de variação calculada pela eq.(5) com dados em R\$ 2003. \* Média geométrica



## BOLSA FAMÍLIA OU DESEMPENHO DA ECONOMIA? DETERMINANTES DA REELEIÇÃO DE LULA EM 2006

MAURICIO CANÊDO-PINHEIRO \*

### Resumo

Este artigo investiga os determinantes dos resultados das eleições presidenciais brasileiras de 2006, em particular o papel do programa Bolsa Família e do desempenho econômico na migração da base eleitoral de Lula para regiões menos desenvolvidas. Os resultados indicam que, entre eleitores habituais de Lula, a participação no programa não aumenta a probabilidade de voto nesse candidato em 2006. Entre eleitores habituais de outros candidatos, esse efeito é positivo. Além disso, um maior crescimento econômico somente significou votos adicionais para Lula entre os eleitores mais ricos. Por fim, o impacto do programa Bolsa Família se mostrou bastante superior ao do desempenho econômico. No entanto, nenhum dos dois fatores foi capaz de explicar satisfatoriamente toda a extensão da mudança observada em 2006.

**Palavras-chave:** Bolsa Família; Lula; Eleições Presidenciais; Viés de Agregação.

### Abstract

This paper investigates the determinants of the outcome of Brazilian presidential elections in 2006, in particular the role of the Bolsa Família Program and the economic performance in the migration of the electoral base of Lula for less developed regions. Among the usual Lula's electorate, joining the Bolsa Família program did not increase the likelihood of voting in this candidate. Among the other voters, this effect was positive. Moreover, higher economic growth only resulted in more votes for Lula among the richest voters. Regarding the elections results, the impact of Bolsa Família proved to be quite superior than the economic performance. However, neither of the two factors was able to satisfactorily explain the full extent of the observed change in 2006.

**Keywords:** Bolsa Família; Lula; Presidential Elections; Aggregation Bias  
**JEL classification:** H53, I38.

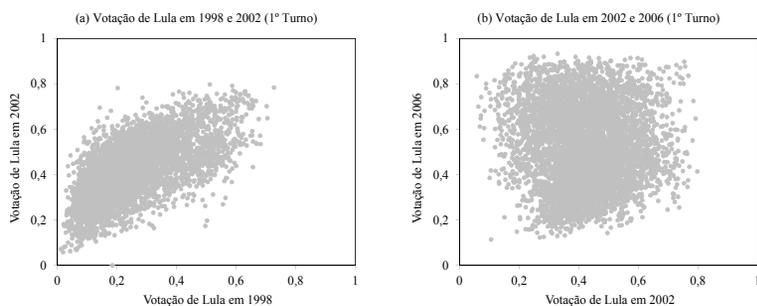
**DOI:** <http://dx.doi.org/10.1590/1413-8050/ea100264>

---

\* BRE/FGV e FCE/UERJ. E-mail: mauricio.pinheiro@fgv.br

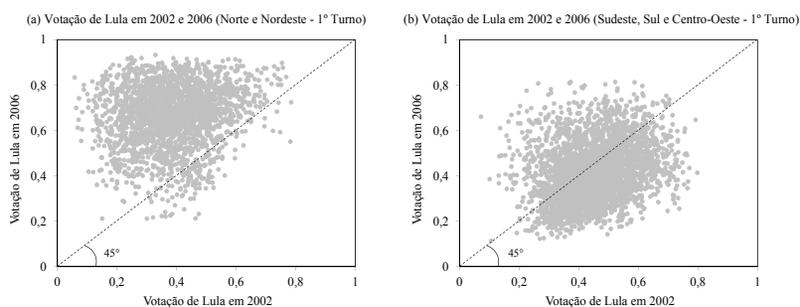
## 1 Introdução

Embora tenha vencido as eleições presidenciais de 2002 e 2006 com praticamente a mesma votação percentual, ficou clara a mudança no perfil da votação do candidato reeleito, Luiz Inácio Lula da Silva (Lula). A Figura 1 ilustra essa afirmação. Cada ponto representa um município brasileiro. A princípio, seria esperada alguma relação positiva entre as votações obtidas por Lula entre duas eleições consecutivas, como ocorreu entre 1998 e 2002. Entretanto, nota-se que praticamente não há relação entre os percentuais de votos obtidos por Lula no primeiro turno, quando comparadas as eleições de 2002 e 2006.<sup>1</sup> Se tradicionalmente Lula era bem sucedido em regiões mais desenvolvidas do Brasil, em 2006 o candidato do Partido dos Trabalhadores (PT) viu sua base eleitoral migrar para regiões de menor desenvolvimento. Essa mudança teve uma contrapartida do ponto de vista geográfico (ver Figura 2). Lula melhorou seu desempenho na vasta maioria dos municípios das regiões Norte e Nordeste (são os pontos localizados acima da reta de 45°). Nas demais regiões a tendência de melhora e piora foi mais equilibrada.



Fonte: Tribunal Superior Eleitoral

**Figura 1:** Votação de Lula em Diferentes Eleições (% dos votos válidos no 1º Turno)



Fonte: Tribunal Superior Eleitoral

**Figura 2:** Votação de Lula em 2002 e 2006 (% dos votos válidos no 1º Turno)

<sup>1</sup>Esse último resultado vale também para a votação no segundo turno. Obviamente, não é possível comparar as votações de Lula no segundo turno em 1998 e 2002, pois em 1998 a eleição foi decidida no primeiro turno.

Muitas hipóteses foram levantadas para explicar tal mudança. Em Hunter & Power (2007) elas foram organizadas em três grandes categorias (não excludentes). Em primeiro lugar, os eleitores mais pobres e menos escolarizados teriam menos acesso à informação sobre os escândalos de corrupção que surgiram no primeiro mandato de Lula. Em segundo lugar, os eleitores de menor renda seriam os mais beneficiados com a política econômica do primeiro governo de Lula. Por último, as políticas sociais, particularmente o programa Bolsa Família, explicariam grande parte do desempenho eleitoral de Lula nas regiões menos desenvolvidas. Além disso, Zucco (2008) aponta evidências de que no Brasil o candidato a presidente ligado ao governo sempre obtém desempenho eleitoral relativamente melhor em regiões menos desenvolvidas, naturalmente mais dependentes do governo. Tal regularidade também ajudaria a explicar o desempenho de Lula nessas regiões.

Sobre a relação entre os programas de transferência de renda e o desempenho de Lula nas eleições de 2006, existe um amplo debate público sobre os dividendos eleitorais do programa Bolsa Família.<sup>2</sup> Esse debate foi transplantado para o ambiente acadêmico. Os resultados são bastante díspares: variando desde evidências de que o programa não foi determinante para a reeleição de Lula (Shikida et al. 2009) até conclusões de que o programa teve papel decisivo na vitória do candidato do PT (Marques et al. 2009). A propósito, essa literatura não é restrita ao caso brasileiro: há artigos que investigam o impacto de programas de transferência de renda nos resultados de eleições em outros países (ver Labonne (2013) e De La O (2013), por exemplo).

A respeito do papel do desempenho da economia brasileira na mudança do padrão eleitoral de Lula em 2006, muitos autores defendem a prevalência de fatores econômicos. Para eles, os ganhos de bem-estar ligados à melhora das condições do mercado de trabalho para a parcela mais pobre da população e ao impacto do câmbio valorizado e da inflação controlada na cesta de consumo dos estratos de renda mais baixos — e não os ganhos auferidos como os programas de transferência de renda — teriam sido os fatores decisivos nas eleições presidenciais de 2006 (Shikida et al. 2009). Entretanto, apenas artigos mais recentes (ver, por exemplo, Souza & Cribari Neto (2013)) tentaram inferir diretamente o efeito das variáveis econômicas na reeleição de Lula e contrapô-las ao efeito do programa Bolsa Família. Por exemplo, Shikida et al. (2009) inferem a importância dos fatores econômicos de maneira indireta: como o efeito do programa Bolsa Família mostra-se pequeno, o resultado das eleições é creditado ao desempenho da economia.<sup>3</sup> Sendo assim, outra contribuição deste artigo é a inclusão de uma medida de desempenho econômico na análise quantitativa do desempenho eleitoral de Lula em 2006, o que possibilita a comparação direta da importância dos fatores econômicos e do programa Bolsa Família.

Em termos metodológicos, a literatura sobre o tema costuma lançar mão de modelos lineares. Esses modelos não são adequados para descrever variáveis que assumem a forma de proporções. Nesse caso, são necessários modelos que se ajustem ao fato de que a variável dependente — proporção de votos em Lula — somente pode assumir valores entre zero e um. Nesse sentido, este trabalho busca suprir essa lacuna ao considerar adequadamente essa questão.

---

<sup>2</sup>Para um resumo deste debate na imprensa ver Marques et al. (2009).

<sup>3</sup>Na verdade, os próprios autores reconhecem a necessidade de análise quantitativa adicional para identificar e medir adequadamente a importância relativa de cada fator.

No entanto, existe outro problema metodológico, muito mais relevante, ainda não tratado nessa literatura. Por conta do mecanismo de voto secreto, é necessário o uso de dados agregados para investigar o impacto do programa Bolsa Família na votação de Lula em 2006. Tal procedimento é suscetível ao que se convencionou denominar falácia ecológica (King 1997) ou viés de agregação (Stoker 1993). Se os eleitores têm características heterogêneas, não seria possível inferir os padrões de comportamento individual com base em dados agregados.<sup>4</sup> Desse modo, conclusões tiradas a partir de dados agregados são potencialmente viesadas e podem não refletir o padrão de escolha individual. As exceções seriam situações em que o modelo individual é linear, o que claramente não é o caso quando a variável de decisão do eleitor é discreta (votar ou não votar em Lula, por exemplo).<sup>5</sup> Trata-se de problema similar à agregação de demandas individuais, tal como descrito no trabalho seminal de Gorman (1953).

Sendo assim, outra contribuição deste artigo — inédita na literatura que estuda os determinantes da reeleição de Lula — é o tratamento adequado dessa questão. Para tanto será utilizado o método de agregação proposto em Kelejian (1995). Esse método, se não permite inferir os parâmetros individuais de escolha, garante que o modelo agregado seja consistente com o modelo individual, evitando o problema de viés de agregação. Além disso, embora a identificação perfeita dos parâmetros individuais não seja possível, consegue-se fazer algumas inferências sobre o comportamento individual dos eleitores, mesmo com dados agregados.

Além desta introdução, este artigo conta com mais cinco seções. A seção 2 descreve o modelo de escolha dos eleitores, enquanto a seção 3 mostra como construir um modelo agregado compatível com as decisões individuais. Por sua vez, a seção 4 descreve os dados utilizados e a especificação exata do referido modelo. A seção 5 traz os resultados da estimação do modelo agregado, bem como uma discussão dos mesmos e algumas inferências sobre o processo de decisão individual dos eleitores. Seguem breves considerações finais.

## 2 As Escolhas Individuais

Defina-se a utilidade indireta do  $i$ -ésimo eleitor, que reside no município  $s \leq S$ , caso tenha votado em Lula nas eleições presidenciais de 2006:

$$u_{is} = \delta + Z_s' \beta + X_{is}' \gamma + \varepsilon_{is}, \quad (1)$$

em que  $S$  é o total de municípios,  $Z_s$  é vetor com variáveis que assumem o mesmo valor para todos os eleitores de um mesmo município,  $X_{is}$  é vetor com variáveis específicas dos eleitores,  $\delta$  é parâmetro,  $\beta$  e  $\gamma$  são vetores de parâmetros e  $\varepsilon_{is}$  é o termo de erro aleatório. Por sua vez, a utilidade indireta do eleitor, caso não vote em Lula, é normalizada para zero.

<sup>4</sup>Dado o caráter secreto dos votos, este é um problema recorrente na literatura que analisa os resultados de eleições (ver, por exemplo, Brown & Payne 1986, Freedman et al. 1991, Heckelman 1997, 2000).

<sup>5</sup>Alguns autores tentam contornar esse problema usando dados individuais extraídos de pesquisas de opinião (Bohn 2011, Licio et al. 2009, Zucco 2013). Entretanto, muitas dessas pesquisas sequer conseguem reproduzir o percentual de votação obtido pelos candidatos nas eleições de 2006. Desse modo, não há como escapar do uso de dados agregados.

Caso tivéssemos acesso aos votos de cada eleitor, somente seria possível observar a decisão de cada um (mas não a utilidade indireta). Sendo assim, é conveniente definir  $D_{is}$  tal que:

$$\begin{aligned} D_{is} &= 1 \text{ se o eleitor escolheu Lula, ou seja, se } u_{is} = \delta + Z'_s\beta + X'_{is}\gamma + \varepsilon_{is} \geq 0 \\ D_{is} &= 0 \text{ caso contrário, ou seja, se } u_{is} = \delta + Z'_s\beta + X'_{is}\gamma + \varepsilon_{is} < 0 \end{aligned} \quad (2)$$

Desse modo, se  $\varepsilon_{is}$  é idêntico e independentemente distribuído com densidade simétrica e distribuição acumulada  $F(Z_s, X_{is})$ , a probabilidade de um indivíduo votar em Lula em determinada eleição, ou seja,  $E(D_{is} | Z_s, X_{is})$ , é dada por:

$$\begin{aligned} \mu_{is} \equiv E(D_{is} | Z_s, X_{is}) &= Pr(D_{is} = 1 | Z_s, X_{is}) \\ &= Pr(\delta + Z'_s\beta + X'_{is}\gamma + \varepsilon_{is} \geq 0) \\ &= 1 - F(-\delta - Z'_s\beta - X'_{is}\gamma) \\ &= F(\delta + Z'_s\beta + X'_{is}\gamma) \end{aligned} \quad (3)$$

Percebe-se que a utilização de modelos lineares se mostra inadequada, pois nesse caso não há como garantir que  $0 \leq \delta + Z'_s\beta + X'_{is}\gamma \leq 1$ , como convém para uma probabilidade. Sob a hipótese de que  $\mu_{is} = F(Z_s, X_{is}) = \frac{e^{\delta + Z'_s\beta + X'_{is}\gamma}}{1 + e^{\delta + Z'_s\beta + X'_{is}\gamma}}$  chega-se aos modelos da família *logit*. Desse modo, a escolha dos eleitores que vivem em um determinado município pode ser modelada como (Train 2003):

$$\ln\left(\frac{\mu_{is}}{1 - \mu_{is}}\right) = \delta + Z'_s\beta + X'_{is}\gamma \quad (4)$$

Os coeficientes de 4 não são diretamente interpretáveis, mas é possível calcular os efeitos marginais de cada variável (Train 2003):

$$\frac{\partial \mu_{is}}{\partial W_{is}} = \mu_{is}(1 - \mu_{is})\vartheta = \frac{e^{\delta + W'_{is}\vartheta}}{(1 + e^{\delta + W'_{is}\vartheta})^2} \quad (5)$$

em que  $\vartheta = \begin{bmatrix} \beta \\ \gamma \end{bmatrix}$  é vetor de parâmetros e  $W_{is} = \begin{bmatrix} Z_s \\ X_{is} \end{bmatrix}$  é vetor que variáveis explicativas.

### 3 Das Escolhas Individuais Para o Resultado Agregado

A estimação dos coeficientes da equação 4 requer informações a respeito de cada eleitor. Infelizmente, uma vez que o voto é secreto, as informações individuais não estão disponíveis. Na estimação de modelos lineares de escolha individual, este problema pode ser contornado pela agregação dos atributos e escolhas individuais em diferentes regiões geográficas (municípios, no caso em análise) (Stoker 1993). Os parâmetros do modelo de decisão individual são então estimados pela regressão das médias da variável dependente nas regiões nas mesmas médias dos regressores e em variáveis *dummy* para as coortes.

Sendo assim, uma questão se impõe: seria possível uma abordagem análoga no caso de modelos da família *logit*, intrinsecamente não lineares? Para responder a essa questão definam-se:

$$\overline{D}_s = \frac{\sum_{i=1}^{N_s} D_{is}}{N_{st}} \quad (6)$$

$$\overline{X}_s = \frac{\sum_{i=1}^{N_s} X_{is}}{N_s}$$

em que  $N_s$  é o número de eleitores da região  $s$  na eleição presidencial de 2006. Perceba que as variáveis de caráter individual passam a ser expressas em termos de proporções ou médias relativas à região. Por exemplo, a variável *dummy* que indica a escolha individual de votar ou não votar em Lula é transformada na proporção de votos desse candidato na região ( $\overline{D}_s$ ). Se a variável for, por exemplo, renda, passa a ser expressa como renda média dos eleitores da região. Tem-se que de forma geral:

$$\eta(\mu_s) = \kappa(Z_s, \overline{X}_s) \quad (7)$$

em que  $\eta(\cdot)$  e  $\kappa(\cdot)$  são funções.

A especificação 7 levanta três questões. A primeira diz respeito à função  $\eta(\cdot)$ : é preciso preservar no modelo agregado a estrutura não linear do modelo de decisão individual descrito em 4 (Stoker 1993)<sup>6</sup>. De outra forma a variável dependente poderá assumir qualquer valor em real quando na verdade é definida no intervalo  $[0, 1]$ . Além disso, modelos lineares resultam em efeitos marginais que independem dos valores das variáveis explicativas. É razoável imaginar que tais efeitos sejam diferentes para valores distintos dessas variáveis.<sup>7</sup> Desse modo, não é adequado o uso de modelos lineares tal como em Nicolau & Peixoto (2007), Soares & Terron (2008), Zucco (2008), Shikida et al. (2009) e Marques et al. (2009).<sup>8</sup> Nesse caso, a escolha mais óbvia é acompanhar o modelo de decisão individual e usar a função logística, ou seja,  $\ln\left(\frac{\mu_s}{1-\mu_s}\right)$ . Aliás, essa escolha é amparada pela constatação de que modelos com função de ligação logística obtêm desempenho bastante superior aos modelos com funções de ligação linear (ver Kieschnick & McCullough (2003) para evidências e referências).

A segunda questão se refere à função  $\kappa(\cdot)$ . Usualmente ela é definida como sendo linear, ou seja,  $\kappa(Z_s, \overline{X}_s) = \delta + Z_s' \beta + \overline{X}_s' \gamma$ . No entanto, tal especificação de  $\kappa(\cdot)$  não é adequada se há heterogeneidade dos indivíduos dentro das regiões (Kelejian 1995). Na verdade, essa especificação parte do pressuposto equivocado de que  $E[\kappa(Z_s, \overline{X}_s)] = \kappa[Z_s, E(X_{is})]$  dentro de cada região (Heckelman & Sullivan 2002). Essa igualdade somente vale no caso em que os atributos individuais não variam entre os eleitores da mesma região. Nesse caso, a princípio, a inclusão dos atributos individuais no modelo implica ter que fazer hipóteses mais específicas sobre a distribuição dos mesmos entre os indivíduos. Obviamente, isso tende a complicar a especificação do modelo agregado, pois se

<sup>6</sup>A função  $\eta(\cdot)$  é usualmente denominada função de ligação (Kieschnick & McCullough 2003).

<sup>7</sup>Para os valores médios das variáveis explicativas é bastante provável que os efeitos marginais sejam semelhantes para modelos não lineares e lineares, uma vez que estes últimos podem ser considerados aproximações de primeira ordem dos demais. No entanto, quanto mais distante das médias, pior será esta aproximação.

<sup>8</sup>Uma exceção é Souza & Cribari Neto (2013), que usam a função de ligação logística.

torna necessário incluir a estrutura da distribuição dos atributos individuais como parte do modelo (Stoker 1993). Esse tema será retomado com mais detalhes na seção 3.1.

Por sua vez, a terceira questão diz respeito à heterocedasticidade dos erros e à distribuição condicional de  $\bar{D}_s$ , tema que será tratado em maior profundidade na seção 3.2.

### 3.1 Correta Agregação das Decisões Individuais

A este respeito, Kelejian (1995) propõe um método que prescinde de qualquer hipótese sobre a distribuição exata dos atributos individuais e, ao mesmo tempo, leva em consideração a heterogeneidade dos mesmos. Esse método é baseado no resultado de que, sob condições bastante gerais, o equivalente agregado do modelo de decisão individual descrito em 4 é:<sup>9</sup>

$$\ln\left(\frac{\bar{D}_s}{1-\bar{D}_s}\right) = \delta + Z'_s\beta + \bar{X}'_s\gamma + g(Z'_s\beta + \bar{X}'_s\gamma) + v_s \quad (8)$$

em que  $g(\cdot)$  é uma função que pode ser interpretada como o viés de agregação causado pela heterogeneidade dos atributos individuais (Kelejian 1995). Em outras palavras, tem-se que  $\kappa(Z'_s, \bar{X}'_s) = \delta + Z'_s\beta + \bar{X}'_s\gamma + g(Z'_s\beta + \bar{X}'_s\gamma)$ . Tal como sugerido em Stoker (1993), o viés e a forma funcional exata de  $g(\cdot)$  dependem da distribuição de  $X_{is}$  entre os eleitores. No entanto, Kelejian (1995) sugere que  $g(\cdot)$  é não linear e pode ser aproximada por um polinômio de ordem  $K$ :

$$g(Z'_s\beta + \bar{X}'_s\gamma) \approx \sum_{k=0}^K (Z'_s\beta + \bar{X}'_s\gamma)^k b_k \quad (9)$$

em que  $b_k$  é o parâmetro associado a  $k$ -ésima potência do polinômio.

Defina-se o vetor  $\bar{W}_s = \begin{bmatrix} Z'_s \\ \bar{X}'_s \end{bmatrix}$ . Substituindo-se 9 em 8 e ignorando-se os erros de aproximação, chega-se a (Kelejian 1995):

$$\underbrace{\ln\left(\frac{\bar{D}_s}{1-\bar{D}_s}\right)}_{\bar{d}_s} = \delta + \bar{W}'_s\vartheta + \sum_{k=0}^K (\bar{W}'_s\vartheta)^k b_k + v_s \quad (10)$$

$$= \underbrace{b_0 + \delta}_a + \underbrace{\bar{W}'_s\vartheta(1 + b_1)}_\lambda + \sum_{k=2}^K (\bar{W}'_s\vartheta)^k \underbrace{\frac{b_k}{(1 + b_k)^k}}_{\phi_k} + v_s$$

Desse modo, em vista das hipóteses feitas, a equação 10 representa a agregação apropriada do modelo exposto em 4. Além disso, a equação 10 pode ser usada para se testar a presença de viés de agregação: Kelejian (1995) sugere que nesse caso a hipótese nula seja definida como  $\phi_2 = \dots = \phi_k = 0$ . Observe que esse teste não é capaz de detectar viés de agregação quando este assume a forma puramente linear.

<sup>9</sup>O resultado não depende, por exemplo, da independência da distribuição dos atributos individuais dentro de cada município.

Revertendo a transformação logística, também é possível calcular (Kelejian 1995):

$$Pr(D_{is} = 1 | \bar{W}_s) \approx \frac{e^{a + \bar{W}'_s \lambda + \sum_{k=2}^K (\bar{W}'_s \lambda)^k \phi_k}}{1 + e^{a + \bar{W}'_s \lambda + \sum_{k=2}^K (\bar{W}'_s \lambda)^k \phi_k}} \equiv \Lambda(\bar{W}'_s, K). \quad (11)$$

Basicamente 11 representa uma média ponderada da probabilidade individual, na qual os pesos são determinados pelas informações regionais agregadas. Note-se que não se trata da previsão de  $\mu_{is}$ , que é a probabilidade condicionada ao vetor  $W_{is}$ . Como para cada eleitor a probabilidade de votar em Lula é função das características individuais, a melhor estimativa para uma pessoa escolhida aleatoriamente em determinada região é a média ponderada de todas as possíveis probabilidades heterogêneas naquela mesma região (Heckelman 1997).

Definindo-se  $\phi_1 = 1$ , não é difícil mostrar que o efeito marginal das variáveis explicativas é:

$$\frac{\partial \bar{D}_s}{\partial \bar{W}_s} = [\Lambda(\bar{W}_s, K)] [1 - \Lambda(\bar{W}_s, K)] \left[ \sum_{k=1}^K (\bar{W}'_s \lambda)^{k-1} k \phi_k \right] \lambda. \quad (12)$$

### 3.2 Heterocedasticidade dos Erros e Distribuição Condicional de $\bar{D}_{st}$

Tomando-se o modelo definido por 10, tem-se que  $\sigma_s \equiv Var(v_s) = \frac{1}{N_s \mu_s (1 - \mu_s)}$  (Maddala 1983, p. 29-30). Além disso, tais variâncias podem ser estimadas pela substituição de por sua contrapartida amostral. Sendo assim, o modelo 10 pode ser estimado por Mínimos Quadrados Não Lineares Generalizados (MQNLG), em que cada observação é ponderada por  $w_s = N_s \bar{D}_s (1 - \bar{D}_s)$ .<sup>10</sup>

Note-se que a variância de  $\bar{D}_s$  deve ser heterocedástica e se aproximar de zero nos limites inferior e superior do intervalo  $[0, 1]$  (Kieschnick & McCullough 2003). Não é difícil mostrar que essa propriedade é satisfeita pelo modelo 10.<sup>11</sup>

Além disso, embora usualmente o modelo 10 seja estimado pela minimização da soma dos quadrados dos resíduos, implicitamente faz-se a hipótese de que  $v \sim N(0, \sigma_s)$  (Kieschnick & McCullough 2003). Em outras palavras, supõe-se que  $\bar{D}_s$  tem distribuição normal.<sup>12</sup> No entanto, essa não é a única possibilidade. Em Kieschnick & McCullough (2003) são abordadas outras opções, como as distribuições beta e simplex.

<sup>10</sup>Na verdade, os pesos devem ser redefinidos para que sua soma seja igual ao número de observações.

<sup>11</sup>Basta notar que  $v \cong \frac{\bar{D}_s - \mu_s}{1 - \mu_s}$  (Maddala 1983, p. 30). Então  $\sigma_s = \left[ \frac{1}{\mu_s (1 - \mu_s)} \right]^2 Var(\bar{D}_s)$ , o que estabelece o resultado de que  $Var(\bar{D}_s) = \frac{\mu_s (1 - \mu_s)}{N_s}$ . Note que  $Var(\bar{D}_s)$  se aproxima de zero quando  $\mu_s$  se aproxima de zero ou um.

<sup>12</sup>Esta hipótese implica assumir que  $\bar{D}_s$  tem distribuição aditiva logística normal.

## 4 Dados e Especificação do Modelo

### 4.1 Dados

A Tabela 1 traz uma lista e uma breve descrição das variáveis utilizadas. As fontes desses dados são basicamente Zucco (2008) e Tribunal Superior Eleitoral (TSE), para as variáveis eleitorais, o Instituto Brasileiro de Geografia e Estatística (IBGE) para as variáveis geográficas, demográficas e socioeconômicas e Ministério do Desenvolvimento Social (MDS) para informações sobre a cobertura do programa Bolsa Família e do número de famílias pobres.

Por sua vez, a Tabela 2 traz algumas estatísticas descritivas dessas variáveis. Fica bastante evidente a disparidade das características socioeconômicas e demográficas entre os municípios da amostra. A título de exemplo, a menor renda per capita da amostra é de R\$ 28,38 por mês e a maior de R\$ 954,65, mais do que trinta vezes superior. Os municípios mais desenvolvidos chegam a ter taxas de analfabetismos inferiores a 1%, enquanto os menos desenvolvidos podem ter mais do que 60% de sua população acima de 15 anos analfabeta.

### 4.2 Especificação do Modelo

Basicamente, foram estimadas diferentes versões do modelo definido em 10, tanto para os votos de Lula no primeiro turno (LULA20061T) quanto para os votos desse candidato no segundo turno (LULA20062T), ambos transformados pela função logística, obviamente. As demais variáveis listadas na Tabela 1 foram usadas como regressores. Adicionalmente, foram incluídas *dummies* estaduais para capturar o efeito de atributos não observáveis e de variáveis omitidas dentro das unidades federativas.<sup>13</sup> As variáveis de maior interesse são BOLSA e CRESC, que capturam o alcance do programa Bolsa Família (BOLSA) e o desempenho econômico no primeiro mandato de Lula em termos de crescimento do PIB per capita (CRESC). As demais variáveis explicativas podem ser agrupadas em quatro categorias: (i) geográficas — URB, DENS, DIST; (ii) políticas e eleitorais — LULA20021T (ou LULA20022T), GOVPT, PREFPT e PT2004; (iii) características socioeconômicas dos eleitores — RENDA, GINI, ANALF, MORT, POBRES, PENT, NBRANCO; (iv) setor público — TXLOCAL, PUBLICO.

## 5 Resultados

A Tabela 3 e a Tabela 4 resumem os resultados encontrados a partir da estimação de diferentes versões do modelo descrito na seção 4.2. Em todas as especificações foi utilizado o método de Mínimos Quadrados Não-Lineares Generalizados (MQNLG). Entretanto, esses resultados podem ser mais bem interpretados se expressos em termos de efeitos marginais. Nesse sentido, a Tabela 5 apresenta os efeitos marginais e elasticidades calculados para algumas das especificações.

<sup>13</sup>Na verdade, para evitar multicolinearidade perfeita com outras variáveis, foram omitidas as *dummies* correspondentes aos estados de Roraima e Acre.

Tabela 1: Variáveis Utilizadas<sup>a</sup>

Variável	Definição	Fonte
LULA20061T	% de votos de Lula sobre votos válidos no 1º turno da eleição de 2006	TSE <sup>b</sup>
LULA20062T	% de votos de Lula sobre votos válidos no 2º turno da eleição de 2006	TSE <sup>b</sup>
BOLSA	% das famílias atendidas pelo programa Bolsa Família em julho de 2006	MDS <sup>b</sup>
CRESC	Taxa de crescimento anual média do PIB per capita entre 2002 e 2006 (%)	IBGE
URB	% da população vivendo em áreas urbanas em 2000	IBGE
DENS	Densidade populacional em 2000 (habitantes por Km <sup>2</sup> )	IBGE
DIST	Distância da capital do estado (Km)	IBGE
LULA20021T	% de votos de Lula sobre votos válidos no 1º turno da eleição de 2002	TSE <sup>b</sup>
LULA20022T	% de votos de Lula sobre votos válidos no 2º turno da eleição de 2002	TSE <sup>b</sup>
GOVPT	Variável binária: 1 se governador do PT em 2006, zero caso contrário	Zucco (2008)
PREFPT <sup>c</sup>	Variável binária: 1 se prefeito do PT em 2006, zero caso contrário	Zucco (2008)
PT2004 <sup>c</sup>	% de votos dos candidatos do PT a prefeito nas eleições municipais de 2004	Zucco (2008)
RENDA	Renda per capita mensal em 2000 (R\$ de 2000)	IBGE
GINI	Índice de desigualdade de Gini em 2000	IBGE
ANALF	% de analfabetos entre pessoas acima de 15 anos em 2000	IBGE
MORT	% de crianças mortas até 5 anos de idade em 2000	IBGE
POBRES	% de famílias pobres – renda familiar per capita inferior a R\$ 100	MDS
PENT	% da população de orientação religiosa pentecostal em 2000	IBGE
NBRANCO	% da população que não declarou ser branca em 2000	IBGE <sup>b</sup>
TXLOCAL	Impostos locais como proporção da receita operacional em 2006 (%)	IBGE
PUBLICO	Administração Pública como proporção do PIB em 2006 (%)	IBGE

<sup>a</sup> Ressalte-se que as variáveis são avaliadas em nível municipal. Por exemplo, a variável LULA20061T se refere à votação de Lula no primeiro turno em cada um dos municípios brasileiros.

<sup>b</sup> Refere-se à fonte primária das informações. Dados retirados de Zucco (2008).

<sup>c</sup> Para Brasília foram calculadas as contrapartidas distritais dessas variáveis, ou seja, com relação ao governador (e não ao prefeito).

Tabela 2: Algumas Estatísticas Descritivas

Variável	Média	Desvio-Padrão	Mínimo	Máximo
LULA20061T	0,510	0,173	0,132	0,934
LULA20062T	0,613	0,164	0,196	0,963
BOLSA	0,289	0,166	0,013	1,000
CRESC	0,014	0,086	-0,452	0,763
URB	0,605	0,238	0,000	1,000
DENS	123,14	625,97	0,100	12881
DIST	250,94	165,76	0,000	1474
LULA20021T	0,437	0,119	0,058	0,798
LULA20022T	0,548	0,123	0,125	0,906
GOVPT	0,065	0,247	0,000	1,000
PREFPT	0,090	0,286	0,000	1,000
RENDA	180,28	100,28	28,38	954,65
GINI	0,562	0,059	0,370	0,820
ANALF	0,207	0,124	0,009	0,607
MORT	0,042	0,029	0,006	0,135
POBRES	0,277	0,157	0,024	0,978
PENT	0,084	0,055	0,001	0,461
NBRANCO	0,450	0,251	0,003	0,987
TXLOCAL	0,053	0,053	0,001	0,510
PUBLICO	0,260	0,140	0,010	0,785

Estatísticas calculadas para uma amostra de 3.398 municípios e não ponderadas pela população. Foram eliminadas as observações em que não estava disponível informação para alguma variável.

### 5.1 Viés de Agregação

Pela inspeção da Tabela 3 e da Tabela 4 percebe-se que o viés de agregação é detectado em todos os modelos (com exceção dos Modelos A1 e B1, obviamente): a hipótese nula de que  $\phi_2 = 0$  (Modelos A2 e B2) ou  $\phi_2 = \phi_3 = 0$  (Modelos A3 e B3) foi fortemente rejeitada.<sup>14</sup>

Note-se que as evidências de Heckelman & Sullivan (2002) mostram que: (i) a presença de viés de agregação aumenta com a população das regiões; (ii) a capacidade do teste proposto por Kelejian (1995) em detectar esse viés também aumenta com o tamanho da população. Nesse sentido, a constatação de que o viés de agregação é importante mesmo com dados municipais lança muitas dúvidas sobre evidências baseadas em dados agregados estaduais, sujeitos a um viés ainda maior. Outro resultado dos mesmos autores indica que o poder desse teste aumenta sensivelmente com a ordem do polinômio ( $K$ ) utilizado na aproximação da função  $g(\cdot)$  e pode ser bastante baixo para valores pequenos de  $K$ . O fato de que a detecção foi conseguida mesmo para  $K = 2$  reforça as evidências de que o viés de agregação não pode ser desprezado. Tanto é verdade, que os efeitos marginais e as elasticidades calculadas para modelos com  $K = 3$  diferem substancialmente dos calculados para modelos com  $K = 1$  (ver mais detalhes no apêndice ao final do artigo).

Também foram estimadas versões do modelo para  $K > 3$ . No entanto, optou-se por apresentar somente os resultados para  $K \leq 3$ , pois as estimati-

<sup>14</sup>Os modelos indexados pela letra A se referem aos votos do 1º turno, os indexados pela letra B dizem respeito ao 2º turno.

vas são bastante similares. Além disso, para  $K > 3$ , o  $R^2$  ajustado e o critério de informação de Akaike indicam que os ganhos em termos de ajuste não compensam as perdas de graus de liberdade.<sup>15</sup> Sendo assim, os Modelos A3 e B3 foram selecionados como preferidos.

## 5.2 Interpretação dos Resultados: Primeiras Impressões

Primeiramente, ressalte-se que os resultados para o primeiro e segundo turnos são semelhantes. Para a maioria das variáveis observou-se o mesmo sinal e efeitos marginais (ou elasticidades) similares nos dois casos. Uma diferença importante é o ganho relativo de relevância da variável CRESC no segundo turno (o efeito marginal quase dobrou). Esse aspecto será explorado mais adiante.

Da Tabela 5 percebe-se também que, tomadas individualmente, as variáveis *LULA20021T* e *LULA20022T* possuem o maior impacto na votação de Lula nas eleições presidenciais de 2006. Um aumento de um ponto percentual na votação de Lula em 2002 significa incremento de 0,4501(0,5452) ponto percentual nos votos desse candidato no primeiro turno (segundo turno) em 2006. Nesse sentido, embora haja indicação de que o padrão de votos de Lula mudou entre 2002 e 2006, a votação do mesmo ainda se mostrou bastante influenciada pelo seu desempenho eleitoral passado.

O sinal das variáveis sociodemográficas relativas às minorias também foi o esperado. Tudo mais constante, a votação de Lula foi superior em municípios com maior participação de negros, pardos e índios na população (*NBRANCO*), o que seria de se esperar dado o apelo de seu primeiro governo no que tange às desigualdades raciais Zucco (2008). Um aumento de 1% nessa variável gera incremento de 0,2520% (0,1946%) na votação de Lula no primeiro turno (segundo turno) em 2006 (ver Tabela 5). O mesmo com relação aos evangélicos ligados às igrejas pentecostais (*PENT*), embora com uma elasticidade inferior.<sup>16</sup>

Das variáveis políticas, a presença de governador do PT (*GOVPT*) e a votação dos candidatos do PT nas eleições municipais de 2004 (*PT2004*) aparecem com o sinal esperado. No entanto, a presença de prefeito do PT apresentou sinal negativo, contrário do esperado. Trata-se do mesmo resultado encontrado em Zucco (2008). Mas, provavelmente, o sinal dessa variável não indica que os prefeitos do PT atrapalharam o desempenho eleitoral de Lula em seus municípios. A variável pode ser uma medida da força dos eleitores habituais de Lula — movimentos sociais organizados, sindicatos e classe média — que provavelmente elegeram ou ajudaram a eleger os prefeitos do PT 2004. Se esse for o caso, o sinal negativo indica que esses eleitores habituais abandonaram Lula em 2006 (Zucco 2008). De todo o modo, trata-se de efeito incremental pequeno: em média a presença de prefeito do PT reduziu em 2,4 (3,8) pontos percentuais a votação de Lula no primeiro turno (no segundo turno). O efeito do governador mostra-se substancialmente maior: aumento médio de 17,39 (28,04) pontos percentuais no primeiro turno (segundo turno) (ver Tabela 5). Entretanto, ressalte-se que, na prática esse efeito teve pouca influência no resultado final, na medida em que em 2006 o PT somente governava estados com pequeno peso eleitoral (Acre, Piauí e Mato Grosso do Sul).

<sup>15</sup>Além disso, para  $K = 4$  as estimativas tornam-se quase todas não significativas.

<sup>16</sup>O sinal é o contrário do encontrado em Zucco (2008).

Tabela 3: ]

Resultados das Regressões — 1º Turno, Variável Dependente  $\ln[LULA20061T/(1-LULA20061T)]$ 

	Modelo A1 (K = 1)	Modelo A2 (K = 2)	Modelo A3 (K = 3)
BOLSA	0,8713* (0,1212)	1,6284* (0,2674)	3,2808* (0,6646)
CRESC	0,1415*** (0,0814)	0,3081** (0,1463)	0,6709** (0,2999)
URB	0,1583* (0,0420)	0,2675* (0,0805)	0,4709* (0,1703)
DENS	0,0000* (0,0000)	0,0000* (0,0000)	0,0001* (0,0000)
DIST	-0,0001* (0,0000)	-0,0002* (0,0001)	-0,0005* (0,0002)
LULA20021T	2,0238* (0,0624)	3,8796* (0,2324)	7,6562* (1,0508)
GOVPT	0,8288* (0,1264)	1,4384* (0,2204)	2,8146* (0,5550)
PREFPT	-0,0738* (0,0150)	-0,1662* (0,0290)	-0,3430* (0,0719)
PT004	0,0699** (0,0295)	0,1589* (0,0554)	0,2949* (0,1141)
RENDA	-0,0010* (0,0001)	-0,0014* (0,0001)	-0,0025* (0,0004)
GINI	-0,7534* (0,1254)	-1,8040* (0,2475)	-3,7744* (0,6748)
ANALF	0,8204* (0,1663)	1,8433* (0,3562)	3,4557* (0,8085)
MORT	0,5459 (0,6322)	1,3383 (1,3550)	2,1313 (2,7191)
POBRES	0,0182 (0,1575)	0,2918 (0,3189)	0,8039 (0,6452)
PENT	0,0047* (0,0012)	0,0086* (0,0022)	0,0159* (0,0048)
NBRANCO	1,2040* (0,0630)	2,4435* (0,1782)	5,1463* (0,7969)
TXLOCAL	0,5446* (0,1086)	0,8561* (0,1898)	1,4586* (0,3961)
PUBLICO	0,4470* (0,0857)	0,9943* (0,1847)	2,0268* (0,4704)
$\phi_2$	-	-0,0408* (0,0017)	-0,0513* (0,0042)
$\phi_3$	-	-	0,0011* (0,0002)
Teste para Viés de Agregação	-	71,97*	82,36*
Número de Observações	3398	3398	3398
R <sup>2</sup>	0,8398	0,8431	0,8436
R <sup>2</sup> Ajustado	0,8377	0,8411	0,8415
Critério de Seleção de Akaike	451,44	381,48	373,09

Os símbolos \*, \*\* e \*\*\* indicam significância a 1%, 5% e 10%, respectivamente. Omitiu-se o valor da constante e dos efeitos estaduais por economia de espaço. Entre parênteses os desvios-padrão das estimativas, robustas à heterocedasticidade da forma como definida na seção 3.2. O teste para viés de agregação reporta a estatística  $\chi^2$  relativa à significância conjunta de  $\phi_2, \dots, \phi_K$  (mais detalhes ver seção 3.1).

**Tabela 4: ]**Resultados das Regressões — 2º Turno, Variável Dependente  $\ln[LULA20062T/(1-LULA20062T)]$ 

	Modelo B1 (K = 1)	Modelo B2 (K = 2)	Modelo B3 (K = 3)
BOLSA	0,7606* (0,1205)	1,3456* (0,2326)	2,5214* (0,6288)
CRESC	0,3124* (0,0764)	0,5218* (0,1303)	1,0432* (0,3253)
URB	0,1910* (0,0399)	0,2935* (0,0700)	0,5402* (0,1693)
DENS	0,0000* (0,0000)	0,0000* (0,0000)	0,0001* (0,0000)
DIST	-0,0002* (0,0000)	-0,0004* (0,0000)	-0,0009* (0,0002)
LULA20022T	2,5165* (0,5773)	4,3977* (0,2664)	8,4817* (1,6855)
GOVPT	0,8300* (0,1126)	1,3747* (0,1953)	2,6411* (0,6821)
PREFPT	-0,0093* (0,0142)	-0,1737* (0,0266)	-0,3323* (0,0778)
PT2004	0,0262 (0,0280)	0,0454 (0,0479)	0,0835 (0,0932)
RENDA	-0,0007* (0,0001)	-0,0009* (0,0001)	-0,0017* (0,0004)
GINI	-0,7111* (0,1193)	-1,4328* (0,2213)	-2,8363* (0,6878)
ANALF	1,4776* (0,1673)	2,6132* (0,3293)	4,8197* (1,0868)
MORT	0,0832 (0,6269)	-0,7224 (1,1809)	-1,3468 (2,2800)
POBRES	0,0295 (0,1540)	0,2185 (0,2771)	0,5304 (0,5435)
PENT	0,0014 (0,0011)	0,0033*** (0,0019)	0,0061 (0,0039)
NBRANCO	1,1542* (0,0608)	2,1337* (0,1697)	4,2940* (0,9345)
TXLOCAL	0,3713* (0,1011)	0,5134* (0,1675)	0,8656** (0,3478)
PUBLICO	0,4665* (0,0832)	0,8907* (0,1642)	1,7401* (0,4776)
$\phi_2$	-	-0,0327* (0,0012)	-0,0463* (0,0039)
$\phi_3$	-	-	0,0009* (0,0002)
Teste para Viés de Agregação	-	52,66*	58,50*
Número de Observações	3398	3398	3398
R <sup>2</sup>	0,8648	0,8668	0,8671
R <sup>2</sup> Ajustado	0,863	0,8651	0,8653
Critério de Seleção de Akaike	40,86	-9,80	-13,64

Os símbolos \*, \*\* e \*\*\* indicam significância a 1%, 5% e 10%, respectivamente. Omitiu-se o valor da constante e dos efeitos estaduais por economia de espaço. Entre parênteses os desvios-padrão das estimativas, robustas à heterocedasticidade da forma como definida na seção 3.2. O teste para viés de agregação reporta a estatística  $\chi^2$  relativa à significância conjunta de  $\phi_2, \dots, \phi_K$  (mais detalhes ver seção 3.1).

**Tabela 5:** Efeitos Marginais e Elasticidades

	Modelo A3 (K = 3)		Modelo B3 (K = 3)	
	1º Turno		2º Turno	
	Efeito Marginal	Elasticidade	Efeito Marginal	Elasticidade
BOLSA	0,1930* (0,0067)	0,1034* (0,0017)	0,1621* (0,0052)	0,0735* (0,0014)
CRESC	0,0395** (0,0157)	0,0010** (0,0004)	0,0671* (0,0159)	0,0014* (0,0003)
URB	0,0277* (0,0001)	0,0311* (0,0001)	0,0347* (0,0001)	0,0330* (0,0001)
DENS	0,0000* (0,0000)	0,0008* (0,0000)	0,0000* (0,0000)	0,0007* (0,0000)
DIST	-0,0000* (0,0000)	-0,0149* (0,0035)	-0,0001* (0,0000)	-0,0218* (0,0027)
LULA20021T	0,4503* (0,0373)	0,3646* (0,0159)	-	-
LULA20022T	-	-	0,5452* (0,0509)	0,4685* (0,0153)
GOVPT	0,1739* (0,0178)	-	0,2804* (0,0195)	-
PREFPT	-0,0240* (0,0031)	-	-0,0380* (0,0031)	-
PT2004	0,0173* (0,0006)	0,0117* (0,0043)	0,0054 (0,0059)	0,0031 (0,0034)
RENDA	-0,0001* (0,0000)	-0,0488* (0,0058)	-0,0001* (0,0000)	-0,0305* (0,0046)
GINI	-0,2220* (0,0263)	-0,2309* (0,0279)	-0,1823* (0,0258)	-0,1605* (0,0222)
ANALF	0,2032* (0,0004)	0,0778* (0,0002)	0,3098* (0,0004)	0,1003* (0,0001)
MORT	0,1254* (0,0001)	0,0098* (0,0000)	-0,0866* (0,0001)	-0,0057* (0,0000)
POBRES	0,0473 (0,0345)	0,0242 (0,0189)	0,0341 (0,0340)	0,0148 (0,0147)
PENT	0,0009* (0,0002)	0,0146* (0,0040)	0,0004 (0,0002)	0,0052 (0,0031)
NBRANCO	0,3027* (0,0001)	0,2521* (0,0001)	0,2760* (0,0001)	0,1945* (0,0001)
TXLOCAL	0,0858* (0,0207)	0,0085* (0,0022)	0,0556* (0,0208)	0,0047* (0,0017)
PUBLICO	0,1192* (0,0002)	0,0574* (0,0001)	0,1118* (0,0002)	0,0456* (0,0001)

Os símbolos \*,\*\* e \*\*\* indicam significância a 1%, 5% e 10%, respectivamente. Efeitos marginais e elasticidades calculados nas médias amostrais não ponderadas pela população dos municípios (ver Tabela 2). Para as variáveis binárias *GOVPT* e *PREFPT* o valor calculado se refere ao efeito incremental de uma mudança discreta (de zero para um). Os desvios-padrão foram calculados pelo método Delta.

Nota-se também que, a despeito da melhora do desempenho eleitoral de Lula nas regiões menos desenvolvidas, tudo mais constante, o candidato do PT ainda obteve maior votação proporcional em regiões mais urbanizadas (*URB*), mais densamente povoadas (*DENS*), mais próximas das capitais (*DIST*) e menos desiguais (*GINI*).<sup>17</sup> O mesmo pode ser dito a respeito dos municípios menos dependentes de transferências do governo estadual e federal (*TXLOCAL*).<sup>18</sup> Nesse sentido, os resultados indicariam que, a princípio, Lula foi relativamente menos votado nos chamados ‘rincões’. Entretanto, os efeitos marginais com relação à renda per capita (*RENDA*), taxa de analfabetismo (*ANALF*), proporção de famílias pobres (*POBRES*) e do tamanho do setor público na economia (*PUBLICO*) indicam que tal conclusão é precipitada. Os resultados da Tabela 5 indicam que um incremento de 1% na renda per capita municipal gera uma redução de 0,0488% (0,0305%) na votação de Lula no primeiro turno (segundo turno) das eleições presidenciais de 2006. A mesma redução na taxa de analfabetismo (ou incremento na participação da administração pública no PIB municipal) gera aumentos de magnitude semelhantes na variável de interesse. No caso da proporção de famílias pobres o efeito não se mostrou estatisticamente significativo (ver Tabela 5).

### 5.3 Bolsa Família ou Economia?

Na prática, a diferença entre o impacto do programa Bolsa Família e do desempenho econômico nas eleições presidenciais é bem maior do que os efeitos marginais fazem supor. A média (sem ponderação) do crescimento do PIB per capita entre os municípios da amostra no período 2002-2006 é 1,4% (ver Tabela 2). Nesse caso, o aumento de um ponto percentual implica quase dobrar o crescimento médio no período. Por sua vez, o alcance médio do programa Bolsa Família é de 28,9% (ver Tabela 2). Em outras palavras, melhorar em um ponto percentual o desempenho econômico significa elevar em 70% a taxa de crescimento do PIB, enquanto o mesmo aumento no alcance do programa Bolsa Família implica incrementar em apenas 3,5% o número de famílias beneficiadas. Não por acaso, tomando-se as elasticidades da Tabela 5, um incremento de 1% na variável *CRESC* eleva em somente 0,0010% (0,0014%) os votos no candidato do PT no primeiro turno (segundo turno). O mesmo aumento percentual em *BOLSA* resulta em elevação de 0,1033% dos votos de Lula no primeiro turno e 0,0735% no segundo turno.

A diferença dos efeitos do programa Bolsa Família e do desempenho econômico também pode ser compreendida com a ajuda de exercícios contrafactuais. Nesse caso, como os resultados são similares para os dois turnos, o foco será mantido no segundo turno.<sup>19</sup> Tome-se primeiro o programa de transferência de renda. Na Figura 3, cada ponto do painel (a) é o valor previsto da votação de Lula no segundo turno caso o programa Bolsa Família não existisse e se a taxa de crescimento no município fosse nula (calculados a partir do modelo B3 — ver Tabela 4). Esses valores são contrapostos à votação do candidato do PT nas eleições presidenciais de 2002. Além disso, é reportado (acima à esquerda) o resultado da regressão da votação de Lula em 2006 sem

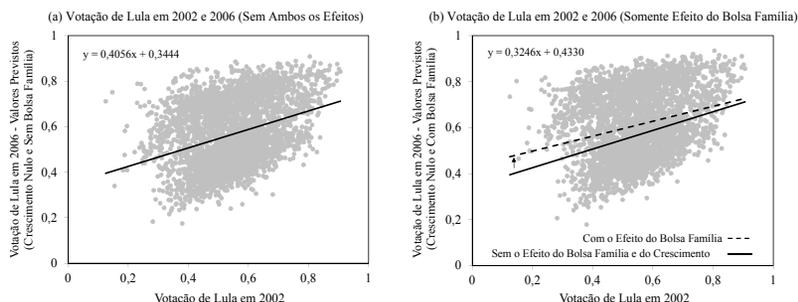
<sup>17</sup>O sinal da variável *GINI* se mostrou contrário ao encontrado em Shikida et al. (2009).

<sup>18</sup>O aumento da taxa de mortalidade infantil (*MORT*) também deprime a votação de Lula, mas apenas no segundo turno. No primeiro turno o sinal encontrado foi positivo.

<sup>19</sup>A diferença principal é que no primeiro turno a relação entre a votação de Lula nas eleições de 2002 e 2006 é mais fraca do que no segundo turno.

os efeitos do crescimento e do programa Bolsa Família (eixo das ordenadas) na votação de Lula em 2002 (eixo das abscissas), bem como a reta ajustada correspondente. O painel (b) reporta as mesmas informações do painel (a), mas considerando a cobertura observada do programa.

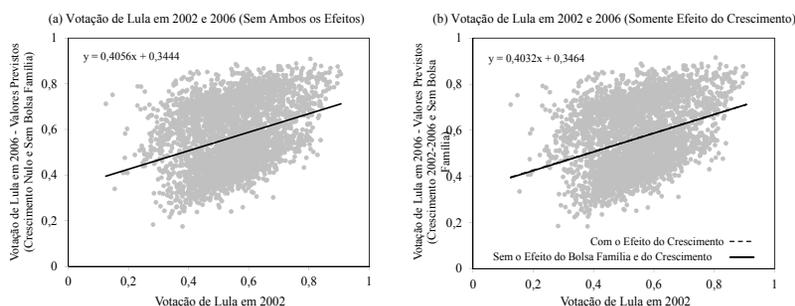
O que se percebe é que a distribuição de benefícios do programa Bolsa Família eleva o intercepto (de 0,3444 para 0,4330) e reduz a inclinação (de 0,4056 para 0,3246) da reta ajustada. Nota-se que, na média, o programa elevou a votação de Lula em todos os municípios, mas que o aumento foi maior justamente naqueles em que o seu desempenho foi relativamente pior em 2002.



Valores calculados a partir do modelo B3 (ver Tabela 4).

**Figura 3:** Votação de Lula em 2006: Efeito do Programa Bolsa Família (2º Turno)

Exercício análogo foi feito para o desempenho da economia, medido pela taxa de crescimento do PIB per capita (ver Figura 4). No entanto, o efeito do desempenho da economia foi bastante modesto se comparado ao impacto do programa Bolsa Família, tanto que sequer é discernível a olho nu como no gráfico da Figura 3.



Valores calculados a partir do modelo B3 (ver Tabela 4).

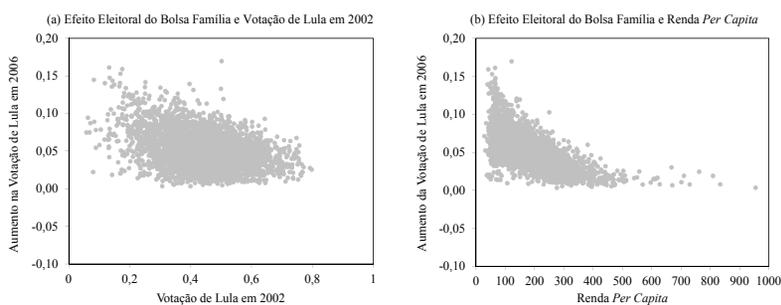
**Figura 4:** Votação de Lula em 2006: Efeito do Crescimento do PIB Per Capita (2º Turno)

A Figura 5 e a Figura 6 ilustram a diferença entre os impactos do programa Bolsa Família e do desempenho econômico. No painel (a) da Figura 5 fica claro que o programa influenciou positivamente o desempenho eleitoral de Lula em 2006, e que esse efeito foi mais importante nos municípios em que tradicionalmente o seu desempenho era pior. A Tabela 6 e a Tabela 7 reforçam esse argumento. O efeito do programa Bolsa Família nos estados das regiões

Norte e Nordeste foi superior ao impacto nos demais estados. Por exemplo, em Alagoas o programa aumentou em 9,45 (7,64) pontos percentuais a votação de Lula no primeiro turno (segundo turno), enquanto no Rio de Janeiro e São Paulo o incremento foi de somente 1,51 (0,98) e 1,92 (1,65) ponto percentual, respectivamente.

O painel (b) da Figura 5 explora a ligação entre o desempenho de Lula nas eleições presidenciais de 2002 e o nível de desenvolvimento dos municípios. O efeito do programa Bolsa Família na votação de Lula foi superior nos municípios mais pobres.

Por sua vez, o painel (a) da Figura 6 indica que o efeito eleitoral do crescimento foi relativamente menos importante do que o impacto do programa Bolsa Família, sendo praticamente nulo na média dos municípios da amostra. Mais uma vez, a Tabela 6 e a Tabela 7 reforçam este argumento: tomando-se o Brasil com um todo, o efeito do programa Bolsa Família na votação de Lula em 2006 foi de 3,30 (2,68) pontos percentuais no primeiro turno (segundo turno), o equivalente a pouco mais de 3,1 milhões (2,5 milhões) de votos. Trata-se de aproximadamente dezessete vezes ( $3,30 \div 0,19$ ) o efeito do crescimento no primeiro turno e mais do que oito vezes ( $2,68 \div 0,31$ ) no segundo turno.<sup>20</sup> É um impacto modesto, principalmente quando se considera que em 2006 o número de famílias atendidas pelo programa era aproximadamente onze milhões. Além disso, o painel (b) da Figura 6 mostra que a relação ténue entre o efeito eleitoral do desempenho da economia com a votação de Lula em 2002 se repete quando se considera a renda *per capita* dos municípios.



**Figura 5:** Efeito Eleitoral do Programa Bolsa Família (2º Turno)

#### 5.4 Nem Bolsa Família nem Economia: o que Mudou Afinal?

Da seção 5.3 percebe-se que o desempenho da economia e o programa Bolsa Família, principalmente este último, tiveram impacto no resultado das eleições presidenciais de 2006. No entanto, fica claro que, sozinhos, não são capazes de explicar a mudança do padrão de votos de Lula entre 2002 e 2006. A Figura 7 ilustra esse argumento para o segundo turno (os resultados para o primeiro turno são similares). A linha pontilhada de 45° indica a votação

<sup>20</sup>Novamente, a diferença entre os impactos do programa Bolsa Família e do crescimento da economia na votação de Lula no primeiro e no segundo turno está em linha com os resultados encontrados na Tabela 5. O efeito marginal da variável *BOLSA* é um pouco menor no segundo turno, enquanto o efeito marginal da variável *CRESC* é quase o dobro.

**Tabela 6:** Votação de Lula: Valores Previstos por Estados — 1º Turno (Modelo A3)

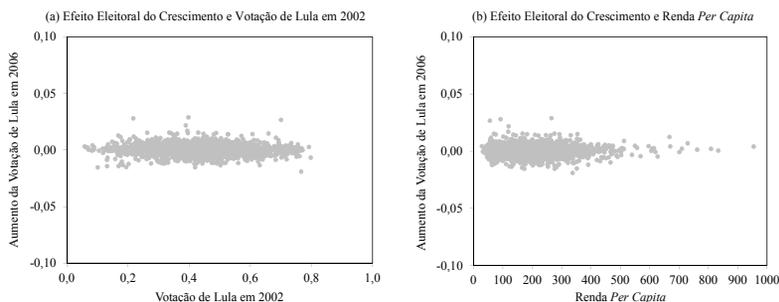
	Observado <sup>a</sup>	Previsto	Previsto		Previsto		Efeito do Bolsa Família	Efeito do Crescimento
			Sem os Efeitos	Ambos	Somente do Bolsa Família	Efeito do Crescimento		
			[A]	[B]	[C]	[B]-[A]		
RO	46,30%	46,60%	41,60%	46,40%	41,80%	4,82%	0,20%	
AC	42,00%	42,60%	35,40%	42,30%	35,70%	6,88%	0,35%	
AM	77,40%	77,10%	74,80%	77,10%	74,80%	2,30%	0,01%	
RR	25,70%	26,10%	19,00%	25,60%	19,50%	6,64%	0,49%	
PA	50,20%	50,10%	45,00%	50,00%	45,20%	4,97%	0,15%	
AP	52,70%	52,60%	49,80%	52,40%	50,00%	2,65%	0,16%	
TO	58,70%	59,10%	53,50%	58,60%	54,00%	5,14%	0,56%	
MA	75,20%	74,70%	69,20%	74,40%	69,60%	5,14%	0,35%	
PI	66,60%	66,10%	59,80%	65,90%	60,00%	6,13%	0,27%	
CE	70,40%	70,10%	65,00%	69,90%	65,10%	4,97%	0,18%	
RN	56,50%	56,00%	50,20%	55,80%	50,40%	5,65%	0,23%	
PB	66,50%	66,20%	60,40%	66,10%	60,50%	5,70%	0,13%	
PE	70,30%	69,90%	65,50%	69,90%	65,50%	4,43%	0,01%	
AL	44,30%	43,80%	34,20%	43,60%	34,30%	9,45%	0,15%	
SE	45,30%	45,40%	38,60%	45,20%	38,80%	6,60%	0,26%	
BA	67,20%	66,90%	62,00%	66,80%	62,20%	4,74%	0,15%	
MG	51,00%	51,10%	47,30%	51,00%	47,40%	3,71%	0,12%	
ES	51,60%	51,70%	47,70%	51,40%	48,00%	3,72%	0,29%	
RJ	48,40%	48,20%	46,50%	48,00%	46,70%	1,51%	0,15%	
SP	36,70%	36,80%	34,60%	36,50%	34,80%	1,92%	0,26%	
PR	37,40%	37,50%	34,00%	37,30%	34,10%	3,37%	0,15%	
SC	33,00%	33,70%	31,60%	33,50%	31,80%	1,88%	0,23%	
RS	33,40%	33,60%	30,60%	33,60%	30,70%	2,92%	0,05%	
MS	35,90%	36,00%	32,00%	35,90%	32,20%	3,85%	0,11%	
MT	38,20%	38,40%	34,30%	38,20%	34,60%	3,89%	0,28%	
GO	38,90%	39,30%	35,60%	39,00%	35,90%	3,38%	0,23%	
DF	37,10%	37,10%	35,00%	36,40%	35,60%	1,45%	0,62%	
Brasil	47,30%	47,30%	43,80%	47,10%	44,00%	3,30%	0,19%	

<sup>a</sup> Os valores observados se referem somente aos 3398 municípios incluídos na amostra e, portanto, diferem dos valores observados para o universo de 5582 municípios.

**Tabela 7:** Votação de Lula: Valores Previstos por Estados — 2º Turno (Modelo B3)

	Observado <sup>a</sup>	Previsto	Previsto		Previsto		Efeito do Bolsa Família	Efeito do Crescimento
			Sem Ambos os Efeitos	Somente do Bolsa Família	Somente do Crescimento	Efeito Família		
			[A]	[B]	[C]	[B]-[A]		
RO	56.80%	56.70%	52.30%	56.40%	52.70%	4.05%	0.34%	
AC	52.00%	52.20%	45.70%	51.60%	46.30%	5.95%	0.60%	
AM	86.90%	86.70%	85.10%	86.60%	85.10%	1.56%	0.02%	
RR	38.20%	38.40%	31.20%	37.40%	32.10%	6.26%	0.97%	
PA	59.60%	59.50%	55.00%	59.20%	55.30%	4.21%	0.24%	
AP	69.40%	69.40%	67.20%	69.10%	67.40%	1.91%	0.23%	
TO	70.20%	69.80%	64.70%	69.00%	65.60%	4.29%	0.92%	
MA	84.70%	84.00%	79.70%	83.50%	80.20%	3.88%	0.53%	
PI	77.00%	76.40%	71.00%	76.00%	71.50%	5.01%	0.43%	
CE	81.90%	81.50%	77.50%	81.20%	77.80%	3.71%	0.27%	
RN	67.10%	66.50%	61.40%	66.20%	61.80%	4.76%	0.37%	
PB	76.20%	75.80%	70.80%	75.60%	71.00%	4.77%	0.21%	
PE	78.10%	77.60%	73.80%	77.60%	73.80%	3.76%	0.02%	
AL	59.40%	58.70%	50.80%	58.50%	51.10%	7.64%	0.25%	
SE	58.60%	58.40%	52.60%	58.00%	53.00%	5.43%	0.42%	
BA	79.50%	79.00%	75.20%	78.80%	75.40%	3.62%	0.22%	
MG	65.60%	65.40%	62.40%	65.30%	62.50%	2.90%	0.18%	
ES	64.30%	64.30%	60.90%	63.80%	61.30%	2.99%	0.47%	
RJ	69.20%	69.00%	67.80%	68.80%	68.10%	0.98%	0.21%	
SP	47.60%	47.50%	45.40%	47.10%	45.90%	1.65%	0.45%	
PR	48.80%	48.80%	45.60%	48.50%	45.90%	2.88%	0.27%	
SC	45.30%	45.60%	43.60%	45.20%	44.00%	1.62%	0.40%	
RS	45.20%	45.30%	42.70%	45.20%	42.80%	2.52%	0.09%	
MS	45.00%	45.00%	41.50%	44.90%	41.70%	3.38%	0.19%	
MT	49.60%	49.70%	45.90%	49.20%	46.40%	3.30%	0.48%	
GO	53.70%	53.80%	50.60%	53.40%	51.00%	2.77%	0.39%	
DF	57.00%	57.00%	54.90%	56.00%	55.90%	1.11%	0.97%	
Brasil	59.80%	59.70%	56.70%	59.40%	57.00%	2.68%	0.31%	

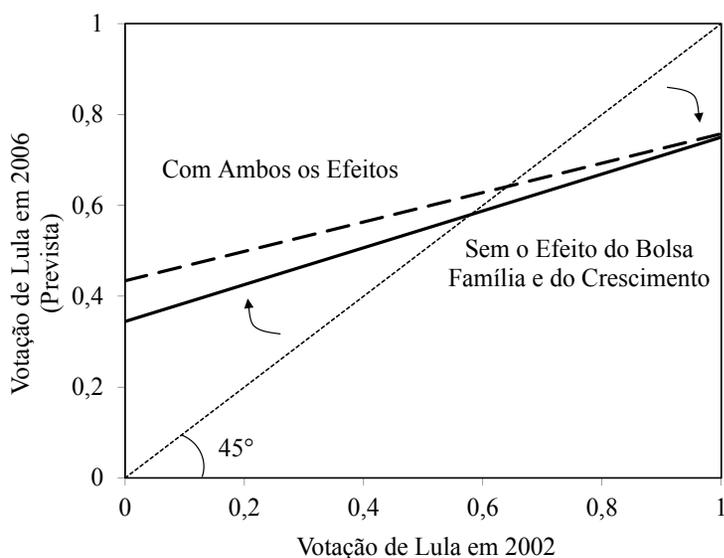
<sup>a</sup> Os valores observados se referem somente aos 3398 municípios incluídos na amostra e, portanto, diferem dos valores observados para o universo de 5582 municípios.



**Figura 6:** Efeito Eleitoral do Desempenho da Economia (2º Turno)

do candidato do PT caso este tivesse mantido o desempenho de 2002. A linha tracejada em negrito aponta a votação média prevista de Lula levando em consideração os efeitos do desempenho econômico e do programa de transferência de renda. Nota-se que ambos somente são capazes de explicar uma parte relativamente pequena da mudança: somente o deslocamento da linha em negrito para a linha tracejada em negrito. A diferença entre a linha de 45° e a linha em negrito — indicada por setas — deve ser creditada a outras variáveis (incluídas ou não no modelo).

Uma possível causa seriam mudanças nas características dos municípios — urbanização, densidade populacional e taxa de analfabetismo, por exemplo — ou no ambiente político — prefeitos e governadores, por exemplo. Entretanto, ou essas variáveis apresentam pequeno efeito marginal na votação de Lula (ver Tabela 5), ou não costumam se alterar significativamente em um período de apenas quatro anos a ponto de explicar tamanha modificação do padrão da votação do candidato do PT.



**Figura 7:** Mudança no Padrão Eleitoral de Lula: Resumo Esquemático (2º Turno)

Nesse sentido, nos municípios mais desenvolvidos, a queda na votação de Lula poderia ser explicada pela percepção dos eleitores mais esclarecidos a respeito dos escândalos de corrupção que surgiram antes das eleições de 2006 (Hunter & Power 2007). No entanto, esse fator não explicaria a melhora no desempenho eleitoral de Lula nos municípios menos desenvolvidos.

Desse modo, talvez a melhora da votação de Lula nos municípios em que tradicionalmente ele obtinha desempenho relativamente pior se deva ao fato de que, em 2006, pela primeira vez ele representava o governo estabelecido. Trata-se do argumento levantado em Zucco (2008) que, independentemente de quem é o candidato, os eleitores de municípios menos desenvolvidos, mais dependentes do Estado, seriam mais propensos a votar nas eleições para presidente a favor do governo. O contrário ocorreria nos municípios mais desenvolvidos. Esse argumento tenderia a ser reforçado pelo fato de que Lula não somente foi o candidato do governo, mas também concorreu à reeleição em 2006. Aliás, as análises de eleições e reeleições devem ser diferentes, pois nestas últimas o voto tende a ter vinculação com o desempenho anterior do candidato (Soares & Terron 2008). Nesse sentido, a mudança no perfil de votos de Lula também pode estar relacionada com a percepção dos eleitores sobre o resultado de outras políticas de seu governo (que não o programa Bolsa Família e o crescimento econômico).

### **5.5 O Problema de Identificação, Especificações Alternativas e Alguns Testes de Robustez**

Além do problema de identificação dos parâmetros do modelo de decisão individual a partir de regressões com dados agregados, que motivou o uso da metodologia proposta em Kelejian (1995), há outro problema de identificação que permeia a literatura que investiga o efeito eleitoral do programa Bolsa Família. Ao contrário de países como o México (ver De La O (2013) para detalhes), não houve programas piloto aleatórios, tampouco descontinuidades que possam ser exploradas na identificação da causalidade entre as variáveis de interesse. Além disso, a elegibilidade e a cobertura do programa são altamente correlacionadas com diversas variáveis socioeconômicas, o que dificulta ainda mais essa identificação (Zucco 2013).

Não há muito que fazer a esse respeito, salvo tentar descartar alguns problemas de identificação por meio de testes de robustez, quando possível. Nesse sentido, um possível problema pode estar associado à orientação ideológica dos eleitores. Se os beneficiários do programa Bolsa Família têm um viés ideológico esquerdista, parte do efeito da variável *BOLSA* na votação de Lula pode ser, na verdade, resultado desse viés. Desse modo, foram estimados modelos similares aos apresentados nas seções anteriores para outros candidatos à presidente nas eleições de 2006 (Geraldo Alckmin, Heloísa Helena e Cristovam Buarque). Os dois últimos são ideologicamente alinhados com a esquerda, como Lula. Caso se encontre algum efeito positivo da variável *BOLSA* na votação dos mesmos, seria uma indicação de que o viés esquerdista dos beneficiários contaminou a estimação do efeito do programa Bolsa Família.

Entretanto, os resultados da Tabela 8 mostra que para todos os candidatos o coeficiente da variável *BOLSA* se mostrou negativo. Embora esse resultado não permita descartar completamente a possibilidade de viés ideológico, trata-se de um indício a esse respeito.

**Tabela 8:** Impacto do Programa Bolsa Família na Votação de Outros Candidatos

	Coeficiente da Variável <i>BOLSA</i>
Geraldo Alckmin (1º Turno)	-1,2445* (0,2281)
Geraldo Alckmin (2º Turno)	-1,0423* (0,1697)
Cristovam Buarque (1º Turno)	-0,1424*** (0,0827)
Heloísa Helena (1º Turno)	-0,7328* (0,1283)

Os símbolos \*, \*\* e \*\*\* indicam significância a 1%, 5% e 10%, respectivamente. Entre parênteses os desvios-padrão das estimativas, robustas à heterocedasticidade da forma como definida na seção 3.2. Coeficientes extraídos de modelos semelhantes aos estimados para votação de Lula (ver Tabela 3 e Tabela 4), mas com variáveis específicas do PT (*LULA20021T*, *LULA20022T*, *GOVPT*, *PREFPT*, *PT2004*) substituídas por variáveis correspondentes ao partido de cada candidato. Todos os modelos estimados com  $K = 3$ .

Outro problema de identificação está relacionado ao padrão espacial dos votos de Lula em 2006. Sendo assim, foram estimadas especificações nas quais a variável *BOLSA* interage com variáveis binárias regionais. Se no Nordeste o efeito for mais forte do que em regiões mais desenvolvidas, é possível que o discurso mais populista de Lula tenha tido papel relevante na mudança do seu padrão de votação. Nesse caso, é possível que o efeito estimado do programa Bolsa Família seja, ao menos em parte, resultado da identificação do público-alvo nordestino com Lula. A Tabela 9 indica que para o segundo turno não parece haver tal viés. As interações não são conjuntamente significativas ao nível de confiança de 95%. Também não é possível discernir, do ponto de vista estatístico, os coeficientes das interações que envolvem a variável binária do Nordeste e a do Sudeste. Entretanto, para o primeiro turno, as interações se mostram conjuntamente significativas e o coeficiente da interação que envolve a variável binária do Nordeste é claramente superior ao do Sudeste. Desse modo, os resultados do primeiro turno devem ser encarados com alguma reserva.

Por fim, convém retomar o resultado de que o efeito do programa Bolsa Família e do crescimento da economia é distinto em municípios com diferentes rendas per capita e votações de Lula em 2002. Sendo assim, foram estimados modelos similares aos modelos A3 e B3 (ver Tabela 3 e Tabela 4), mas adicionando-se seguintes interações: *BOLSAxRENDA*, *BOLSAxLULA20021T* (ou *BOLSAxLULA20022T*, para o segundo turno), *CRESCxRENDA* e *CRESCxLULA20021T* (ou *BOLSAxLULA20022T*, para o segundo turno). A Tabela 10 apresenta os coeficientes das variáveis de interesse (*BOLSA* e *CRESC*) e das interações que as envolvem (os demais coeficientes não sofreram alterações importantes e, por isso, não foram apresentados).<sup>21</sup>

Novamente, os coeficientes estimados são de difícil interpretação. De todo modo, é possível perceber que, nos dois turnos, o efeito do programa Bolsa Família se reduz com o incremento da renda *per capita*. Por sua vez, a pro-

<sup>21</sup> Também foram estimados modelos com apenas algumas das interações. Como os resultados não se alteraram muito, somente as estimativas do modelo com todas as interações foram apresentadas.

**Tabela 9:** Resultados de Especificações Alternativas – Interações com Efeitos Regionais

	Modelo A4 (K = 3) (1º Turno)	Modelo B4 (K = 3) (2º Turno)
Variável Dependente	$\ln[LULA20061T/(1 - LULA20061T)]$	$\ln[LULA20062T/(1 - LULA20062T)]$
BOLSA	0,4814 (1,0772)	1,0035 (0,8932)
CRESC	0,6975** (0,3032)	0,9798* (0,3086)
BOLSA × NORTE	1,5047 (1,3795)	0,2767 (1,0622)
BOLSA × NORDESTE	4,5000* (1,2485)	1,6753*** (0,9123)
BOLSA × SUDESTE	1,6552 (1,0932)	1,7418*** (0,9362)
CRESC × SUL	1,8461 (1,1246)	0,8704 (0,9173)
Coefficientes de Todas as Interações São Nulos	29,90 [0,0000]	8,51 [0,0746]
Coefficiente da Interação do Nordeste Igual ao do Sudeste	17,82 [0,0000]	0,02 [0,8906]

Os símbolos \*, \*\* e \*\*\* indicam significância a 1%, 5% e 10%, respectivamente. Omitiu-se o valor dos demais coeficientes por economia de espaço. Entre parênteses os desvios-padrão das estimativas, robustas à heterocedasticidade da forma como definida na seção 3.2. Entre colchetes o p-valor das estatísticas. O modelo A4 é semelhante ao modelo A3 (ver Tabela 3), com a adição das seguintes variáveis: *BOLSA × NORTE*, *BOLSA × NORDESTE*, *BOLSA × SUDESTE* e *BOLSA × SUL*. O modelo B4 é o análogo com relação ao modelo B3 (ver Tabela 4).

**Tabela 10:** Resultados de Especificações Alternativas – Interações com LULA20021T (ou LULA20022T) e RENDA

	Modelo A5 (K=3) (1º Turno)	Modelo B5 (K=3) (2º Turno)
Variável dependente	$\ln[LULA20061T/(1 - LULA20061T)]$	$\ln[LULA20062T/(1 - LULA20062T)]$
BOLSA	8,4816* (1,5041)	10,6467* (2,6943)
CRESC	-2,1941*** (1,1289)	4,1215* (1,5258)
BOLSA × RENDA	-0,0123* (0,0034)	-0,0025 (0,0024)
BOLSA × LULA20021T	-7,4926* (1,8880)	-
BOLSA × LULA20022T	-	-13,6220* (3,5090)
CRESC × RENDA	0,0060* (0,0018)	0,0055* (0,0119)
CRESC × LULA20021T	1,6264 (2,3113)	-
CRESC × LULA20022T	-	-8,2466* (2,6758)

Os símbolos \*, \*\* e \*\*\* indicam significância a 1%, 5% e 10%, respectivamente. Omitiu-se o valor dos demais coeficientes por economia de espaço. Entre parênteses os desvios-padrão das estimativas, robustas à heterocedasticidade da forma como definida na seção 3.2. O modelo A5 é semelhante ao modelo A3 (ver Tabela 3), com a adição das seguintes variáveis: *BOLSA × NORTE*, *BOLSA × NORDESTE*, *BOLSA × SUDESTE* e *BOLSA × SUL*. O modelo B5 é semelhante ao modelo B3 (ver Tabela 4), com a adição das seguintes variáveis: *BOLSA × NORTE*, *BOLSA × NORDESTE*, *BOLSA × SUDESTE* e *BOLSA × SUL*.

porção de votos em Lula nas eleições de 2002 aumenta (reduz) o impacto do crescimento econômico na votação de Lula em 2006 no primeiro (segundo) turno.

Entretanto, uma das vantagens da abordagem econométrica utilizada é que, embora não seja possível identificar os parâmetros do modelo de decisão individual, consegue-se inferir a importância relativa de cada variável na decisão do eleitor. Os parâmetros do modelo agregado permitem inferir algo sobre o comportamento individual. Mais especificamente, a razão entre cada um dos coeficientes do modelo de decisão individual — vetor  $\vartheta$ , ver equação 4 — e cada um dos coeficientes do modelo agregado — vetor  $\lambda$ , ver equação 10 — é a mesma  $(1 + b_1)$  (Heckelman 1997). Em outras palavras, tem-se que  $\lambda = (1 + b_1)\vartheta$ .

Da equação 12 não é difícil perceber que, para o modelo agregado do primeiro turno, os efeitos marginais das variáveis BOLSA e CRESC na votação de Lula em 2006 ( $\overline{D}_s$ ) serão nulos quando:

$$\begin{aligned}\Omega_{BOLSA} &\equiv \lambda_{BOLSA} + \lambda_1 RENDA + \lambda_2 LULA20021T = 0, \\ \Omega_{CRESC} &\equiv \lambda_{CRESC} + \lambda_3 RENDA + \lambda_4 LULA20021T = 0,\end{aligned}\quad (13)$$

em que  $\lambda_{BOLSA}$  é o coeficiente da variável BOLSA,  $\lambda_{CRESC}$  é o coeficiente da variável CRESC,  $\lambda_1$  é o coeficiente da variável BOLSA  $\times$  RENDA,  $\lambda_2$  é o coeficiente da variável BOLSA  $\times$  LULA20021T,  $\lambda_3$  é o coeficiente da variável CRESC  $\times$  RENDA e  $\lambda_4$  é o coeficiente da variável CRESC  $\times$  LULA20021T. Para o segundo turno, o resultado é análogo.

A título de ilustração, em um município em que a votação de Lula foi 100% em 2002 ( $LULA20021T = 1$ ) a renda per capita necessária para tornar nulo o efeito do programa Bolsa Família na votação de Lula em 2006 é:

$$\overline{RENDA} = -\frac{\lambda_2 + \lambda_{BOLSA}}{\lambda_1} \quad (14)$$

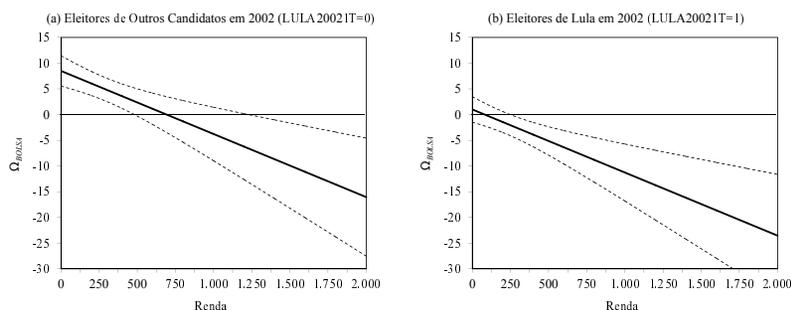
Como  $\lambda = (1 + b_1)\vartheta$ , a condição que vale para o modelo agregado também vale para o modelo de decisão individual. Ou seja, para um eleitor que votou em Lula em 2002, a renda individual que torna nulo o efeito do recebimento do benefício do programa Bolsa Família na probabilidade de se votar nesse mesmo candidato em 2006 é exatamente a mesma calculada em 14:

$$\begin{aligned}RENDA &= \frac{\vartheta_2 + \vartheta_{BOLSA}}{\vartheta_1} \\ &= -\frac{(1 + b_1)\vartheta_2 + (1 + b_1)\vartheta_{BOLSA}}{(1 + b_1)\vartheta_1} \\ &= -\frac{\lambda_2 + \lambda_{BOLSA}}{\lambda_1} \\ &= \overline{RENDA}\end{aligned}\quad (15)$$

em que  $\vartheta_{bolsa}$ ,  $\vartheta_1$  e  $\vartheta_2$  são as contrapartidas individuais de  $\lambda_{BOLSA}$ ,  $\lambda_1$  e  $\lambda_2$ .

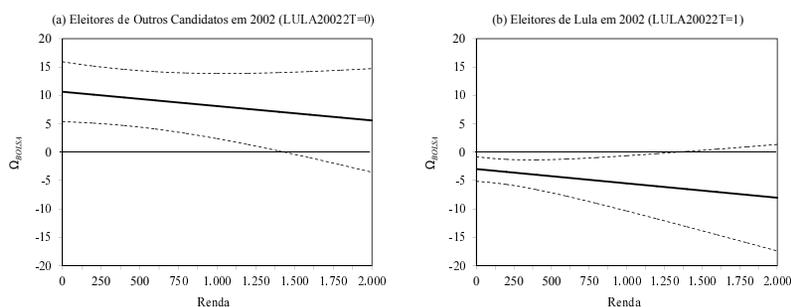
Nesse sentido, a Figura 8 apresenta um exercício bastante informativo para o primeiro turno das eleições. É calculado o valor de  $\Omega_{BOLSA}$  para diferentes níveis de renda individual mensal. Também são discriminados os eleitores habituais de Lula ( $LULA20021T = 1$ ) dos demais eleitores ( $LULA20021T = 0$ ).

O mesmo é feito para o segundo turno na Figura 9. Note-se que, embora o exercício tenha sido feito para valores de renda até R\$ 2.000 mensais, ele somente faz sentido para valores mais baixos, pois somente esses eleitores são elegíveis para participar do programa Bolsa Família.<sup>22</sup> Desse modo, entre os eleitores habituais de outros candidatos que são elegíveis para participar do programa Bolsa Família, o efeito do recebimento do benefício é positivo. Por sua vez, tomando-se o mesmo grupo de renda, mas entre os eleitores habituais de Lula, e considerando-se intervalos de confiança calculados ao nível de significância de 5%, o efeito é nulo (ou algo muito próximo disso).



As linhas pontilhadas indicam o intervalo de confiança calculado a 95%.  
Desvios-padrão calculados pelo método Delta.

**Figura 8:**  $\Omega_{BOLSA}$  e Renda (1º Turno)



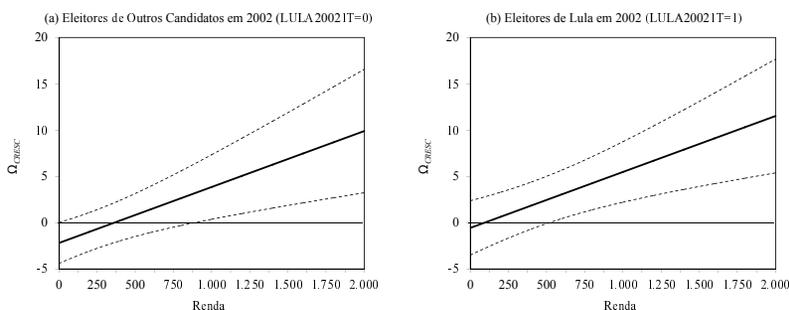
As linhas pontilhadas indicam o intervalo de confiança calculado a 95%.  
Desvios-padrão calculados pelo método Delta.

**Figura 9:**  $\Omega_{BOLSA}$  e Renda (2º Turno)

A Figura 10 (Figura 11) repete o exercício da Figura 8 (Figura 9) para o valor de  $\Omega_{CRESC}$ . No primeiro turno o crescimento econômico somente teve impacto positivo na votação de Lula em 2006 para eleitores mais ricos (a partir de R\$ 890 para eleitores habituais de outros candidatos e a partir de R\$ 520 para eleitores habituais de Lula, levando-se em consideração o intervalo de confiança de 95%). No segundo turno o efeito foi sempre positivo para eleitores habituais de outros candidatos. Para eleitores habituais de Lula o efeito chega a ser negativo entre os eleitores muito pobres (renda abaixo de R\$ 415). Tais resultados parecem indicar que os eleitores mais pobres estão tão alijados

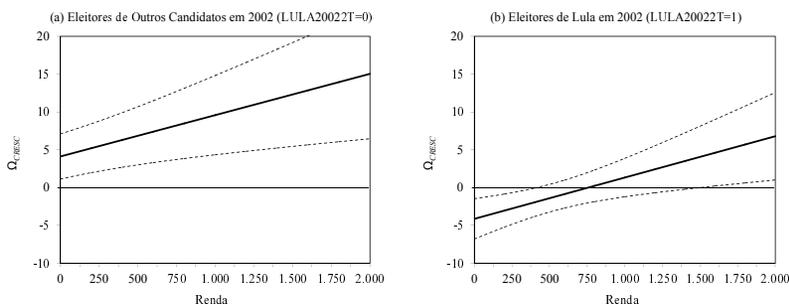
<sup>22</sup>De acordo com dados da PNAD de 2006, 75% das pessoas beneficiadas pelo programa Bolsa Família tinham renda inferior R\$ 575 (a preços de 2000).

dos ganhos advindos do crescimento econômico, que simplesmente não levaram essa variável em consideração quando decidiram seu voto nas eleições de 2006.



As linhas pontilhadas indicam o intervalo de confiança calculado a 95%.  
Desvios-padrão calculados pelo método Delta.

**Figura 10:**  $\Omega_{CRESC}$  e Renda (1º Turno)



As linhas pontilhadas indicam o intervalo de confiança calculado a 95%.  
Desvios-padrão calculados pelo método Delta.

**Figura 11:**  $\Omega_{CRESC}$  e Renda (2º Turno)

## 6 Considerações Finais

O que explicaria a mudança no padrão de votação de Lula nas eleições presidenciais de 2006: desempenho da economia ou a massificação do programa Bolsa Família? Ao contrário de boa parte da literatura anterior, a resposta é: nenhum dos dois. O impacto do programa de transferência de renda se mostrou bastante superior ao do crescimento da economia, mas parece não ser a principal explicação para a migração significativa de votos em direção do candidato do PT nas regiões menos desenvolvidas do Brasil.

Nesse sentido, ganha bastante apelo a conjectura de que, independente do partido e do candidato, os eleitores de regiões menos desenvolvidas, mais dependentes do setor público, seriam mais propensos a votar no governo estabelecido (Zucco 2008). O mesmo pode ser dito com relação à hipótese de Hunter & Power (2007) sobre os efeitos dos escândalos de corrupção na base eleitoral tradicional de Lula. No entanto, essas hipóteses são difíceis de serem empiricamente testadas. Uma possível abordagem é a utilização de dados de várias eleições.

Ressalte-se também que a metodologia empregada nesse artigo (Kelejian 1995) tem a vantagem de lidar adequadamente com o problema de viés de agregação. Além disso, mesmo com dados agregados, essa metodologia permite alguma inferência sobre o comportamento individual dos eleitores. A esse respeito, entre eleitores habituais de Lula, o recebimento do programa Bolsa Família não aumenta a probabilidade de voto nesse candidato em 2006. Entre os eleitores habituais de outros candidatos, esse efeito é positivo. Por sua vez, um maior crescimento econômico somente se traduziu em mais votos para Lula em 2006 entre os eleitores mais ricos. Esses resultados parecem indicar que os eleitores mais pobres estão tão alijados dos ganhos advindos do crescimento econômico, que simplesmente não levaram essa variável em conta quando decidiram seu voto nas eleições de 2006.

Por fim, boa parte das variáveis explicativas foi incluída com valores do ano 2000. Trata-se de uma restrição imposta pela agregação das decisões individuais em nível municipal, cuja maioria das informações remete ao Censo Demográfico de 2000. Sendo assim, não é possível, por exemplo, um exercício contrafactual para a renda *per capita* (RENDA) como os realizados para as variáveis que medem o alcance do programa Bolsa Família e o crescimento econômico. Nesse caso, a única solução é o uso de dados agregados estaduais, mais sujeitos ao viés de agregação, porém com informações disponíveis para os anos de 2002 e 2006.

Nesse sentido, o uso de dados da Pesquisa Nacional de Amostras por Domicílio (PNAD) pode ser bastante promissor, inclusive porque o suplemento de 2006 permite identificar se determinado indivíduo pertence a uma família beneficiária do programa Bolsa Família (e o valor do benefício recebido), o que abriria caminho para o uso de modelos logit com coeficientes aleatórios (Nevo 2000). Esse tipo de modelagem permite considerar atributos individuais não observáveis e também a distribuição dos atributos individuais observáveis dos eleitores.<sup>23</sup> Além disso, com as informações disponíveis na PNAD seria possível identificar melhor o efeito do desempenho da economia, na medida em que, além da taxa de crescimento do PIB, outras variáveis relacionadas ao ambiente econômico — renda do trabalho, desemprego e formalização, por exemplo — poderiam ser usadas. Nesse caso o efeito do crescimento pró-pobre na votação de Lula (Hunter & Power 2007, Shikida et al. 2009) poderia ser mais bem medido. No entanto, tal estratégia somente será adotada em estudos subsequentes.

## Agradecimentos

O autor agradece a César Zucco e Leonardo Monastério pelo compartilhamento das bases de dados usadas respectivamente em Zucco (2008) e Shikida et al. (2009). Agradecimento também a dois pareceristas anônimos, Adriana Hernandez-Perez, Samuel Pessôa, Régis Bonelli e diversos participantes de seminários no IBRE/FGV e na EPGE/FGV pelos comentários. Os erros remanescentes são de responsabilidade do autor e as opiniões expressas não refletem as posições da FGV ou do IBRE.

---

<sup>23</sup>Além disso, ao considerar a distribuição dos atributos individuais, este modelo contorna o problema de viés de agregação. Para uma aplicação deste modelo ao estudo de eleições ver Glasgow (2001).

## Referências Bibliográficas

- Bohn, S. R. (2011), 'Social policy and vote in Brazil. Bolsa Familia and the Shifts in Lula's electoral base', *Latin America Research Review* **46**(1), 54–79.
- Brown, P. J. & Payne, C. D. (1986), 'Aggregate data, ecological regression, and voting transitions', *Journal of American Statistical Association* **81**(394), 452–460.
- De La O, A. (2013), 'Do conditional cash transfers affect electoral behavior? evidence from a randomized experiment in Mexico', *American Journal of Political Science* **57**(1), 1–14.
- Freedman, D. A., Klein, S. P., Sacks, J., Smyth, C. A. & Everett, C. G. (1991), 'Ecological regression and voting rights', *Evaluation Review* **15**(6), 673–711.
- Glasgow, G. (2001), 'Mixed logit models for multiparty elections', *Political Analysis* **9**(2), 116–136.
- Gorman, W. (1953), 'Community preference field', *Econometrica* **21**(1), 63–80.
- Heckelman, J. C. (1997), 'Determining who voted in historical elections: An aggregated logit approach', *Social Science Research* **26**(2), 121–134.
- Heckelman, J. C. (2000), 'Revisiting the relationship between secret ballots and turnout. a new test of two legal-institutional theories', *American Politics Research* **28**(2), 194–215.
- Heckelman, J. C. & Sullivan, T. S. (2002), Testing for aggregation bias in a non-linear framework: Some Monte Carlo results, Mimeo, Winston-Salem, Edwardsville.
- Hunter, W. & Power, T. J. (2007), 'Rewarding Lula: Executive power, social policy, and the Brazilian elections of 2006', *Latin American Politics and Society* **49**(1), 1–30.
- Kelejian, H. H. (1995), 'Aggregated heterogeneous dependent data and the logit model: A suggested approach', *Economics Letters* **47**(3-4), 243–248.
- Kieschnick, R. & McCullough, B. D. (2003), 'Regression analysis of variates observed on (0, 1): percentages, proportions and fractions', *Statistical Modelling* **3**(3), 193–213.
- King, G. (1997), *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregated Data*, Princeton University Press, Princeton.
- Labonne, J. (2013), 'The local electoral impacts of conditional cash transfers: Evidence from a field experiment', *Journal of Development Economics* **104**, 73–88.
- Licio, E. C., Rennó, L. R. & Castro, H. C. O. (2009), 'Bolsa Família e o voto na eleição presidencial de 2006: em busca do elo perdido', *Opinião Pública* **15**(1), 31–54.

- Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.
- Marques, R. M., Leite, M. G., Mendes, A. & Ferreira, M. R. J. (2009), 'Discutindo o papel do programa Bolsa Família na decisão das eleições presidenciais brasileiras de 2006', *Revista de Economia Política* 29(1), 114–132.
- Nevo, A. (2000), 'A practitioner's guide to estimation of random-coefficients logit models of demand', *Journal of Economics & Management Strategy* 9(4), 513–548.
- Nicolau, J. & Peixoto, V. (2007), 'Uma disputa em três tempos: Uma análise das bases municipais das eleições presidenciais de 2006', Anais do XXXI Encontro Nacional da ANPOCS.
- Shikida, C. D., Monasterio, L. M., Araujo Jr, A. F., Carraro, A. & Damé, O. M. (2009), "'It is the economy, companheiro!": an empirical analysis of Lula's re-election based on municipal data', *Economics Bulletin* 29(2), 977–992.
- Soares, G. A. D. & Terron, S. L. (2008), 'Dois Lulas: a geografia eleitoral da reeleição (explorando conceitos, métodos e técnicas de análise geoespacial)', *Opinião Pública* 14(2), 269–301.
- Souza, T. C. & Cribari Neto, F. (2013), 'Uma estimativa do impacto eleitoral do programa Bolsa-Família', *Revista Brasileira de Biometria* 31(1), 79–103.
- Stoker, T. M. (1993), 'Empirical approaches to the problem of aggregation over individuals', *Journal of Economic Literature* 31(4), 1827–1874.
- Train, K. E. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.
- Zucco, C. (2008), 'The president's 'new' constituency: Lula and the pragmatic vote in Brazil's 2006 presidential elections', *Journal of Latin American Studies* 40(1), 29–39.
- Zucco, C. (2013), 'When payouts pay off: Conditional cash transfers and voting behavior in Brazil 2002–10', *American Journal of Political Science* 57(4), 810–822.

## Apêndice A Efeitos marginais, elasticidades e viés de agregação

Na seção 5.1 foi mencionado que a incorporação da correção para o viés de agregação gerava diferenças não desprezíveis nas estimativas de efeitos marginais e de elasticidades. Nesse sentido, a Tabela A.1 ilustra essa afirmação para as duas principais variáveis explicativas utilizadas (*BOLSA* e *CRESC*).

Para o primeiro turno, comparando-se os modelos A1 e A3, nota-se que o efeito marginal da variável *BOLSA* (*CRESC*) é aproximadamente 11,3% menor (11,7% maior) quando se controla o viés de agregação. Analogamente, para o segundo turno (modelos B1 e B3), o efeito marginal da variável *BOLSA* (*CRESC*) é aproximadamente 9,1% menor (11,0% menor). Resultados similares são encontrados para as elasticidades. A interpretação desses resultados é bastante difícil, na medida em que a magnitude e a direção do viés de agregação dependem da distribuição das variáveis explicativas entre os eleitores (ver seção 3.1).

**Tabela A.1:** Efeitos Marginais e Elasticidades — Comparação entre Modelos

		1º Turno			2º Turno		
		Modelo A1	Modelo A2	Modelo A3	Modelo B1	Modelo B2	Modelo B3
BOLSA	Efeito Marginal	0,2176* (0,0059)	0,2094** (0,0852)	0,1930* (0,0067)	0,1783* (0,0049)	0,1763* (0,0060)	0,1621* (0,0052)
	Elasticidade	0,1217* (0,0021)	0,1130* (0,0022)	0,1034* (0,0017)	0,0826* (0,0015)	0,0800* (0,0014)	0,0735* (0,0014)
CRESC	Efeito Marginal	0,0353*** (0,0191)	0,0396** (0,0177)	0,0395** (0,0157)	0,0732* (0,0179)	0,0684* (0,0168)	0,0671* (0,0159)
	Elasticidade	0,0009*** (0,0005)	0,0010** (0,0005)	0,0010** (0,0004)	0,0016* (0,0004)	0,0015* (0,0004)	0,0014* (0,0003)

Os símbolos \*, \*\* e \*\*\* indicam significância a 1%, 5% e 10%, respectivamente Efeitos marginais e elasticidades calculados nas médias amostrais não ponderadas pela população dos municípios (ver Tabela 2). Os desvios-padrão foram calculados pelo método Delta.

# ANÁLISE DA EFICIÊNCIA TÉCNICA DE PROPRIEDADES CITRÍCOLAS DO ESTADO DE SÃO PAULO

FELIPPE CLEMENTE \*  
MARÍLIA FERNANDES MACIEL GOMES †  
VIVIANI SILVA LÍRIO ‡

## Resumo

Este trabalho teve por objetivo analisar a eficiência técnica das propriedades citrícolas do estado de São Paulo. Aplicou-se a análise envoltória de dados (DEA), para calcular os níveis de eficiência técnica, e uma abordagem econométrica, para conhecer os determinantes da eficiência técnica. Os resultados mostraram que grande parte das propriedades citrícolas atua de forma ineficiente e as variáveis que mais contribuem para aumento da eficiência são “escolaridade” e “tempo como produtor rural”. As funções de produção média e de fronteira mostraram-se diferentes quanto à magnitude dos seus parâmetros, indicando que esse instrumental de análise é muito importante e sua estimação e interpretação, de forma correta, pode facilitar a tomada de decisão e melhorar significativamente a alocação dos recursos.

**Palavras-chave:** Eficiência técnica; Setor citrícola; DEA.

## Abstract

This study aimed at analyzing the technical efficiency of citrus farms in the State of São Paulo. For that, data envelopment analysis was applied (DEA) to calculate the levels of technical efficiency, and an econometric approach was used to know the determinants of technical efficiency. The results showed that the majority of citrus properties of São Paulo operate inefficiently and the variables that contribute most to increasing efficiency are "schooling" and "time as farmers." The average production and frontier functions were different regarding the magnitude of its parameters, indicating that the instrumental analysis is very important and its correct estimation and interpretation, can facilitate decision making and significantly improve the allocation of resources.

**Keywords:** Technical efficiency; The citrus sector; DEA.

**JEL classification:** Q12, C25

**DOI:** <http://dx.doi.org/10.1590/1413-8050/ea104786>

\* Universidade Federal de Viçosa — MG. E-mail: felippe.clemente@ufv.br

† Universidade Federal de Viçosa — MG. E-mail: mfmgomes@ufv.br

‡ Universidade Federal de Viçosa — MG. E-mail: vslirio@ufv.br

## 1 Introdução

A cultura da laranja está presente em todos os estados brasileiros em padrões de produção bastante distintos. Ao mesmo tempo, a produção de laranjas é uma atividade concentrada espacialmente no país, uma vez que 96% da produção provém de apenas seis estados, com destaque para São Paulo, responsável por 78% da produção Agriannual (2010).

De acordo com Neves (2005), no contexto do agronegócio brasileiro, o setor citrícola se destaca, pois envolve mais de 24 mil propriedades rurais e emprega diretamente 11,2% da força de trabalho agrícola do estado de São Paulo e 2,2% do Brasil. No ano de 2010, o Brasil foi o principal produtor de laranja do mundo, com 31% da produção, seguido pelos Estados Unidos e pela União Européia, com 16% e 11%, respectivamente. Nas vendas externas o posto ocupado é equivalente: o país é o principal exportador de suco de laranja concentrado congelado (SLCC), com cerca de 80% do mercado internacional. Adicione-se a esta informação o fato de que as exportações de suco de laranja se mantêm, desde 1994, entre 1,1 e 1,4 milhão de toneladas, gerando mais de um bilhão de dólares de divisas. (Neves 2005).

Com relação à produção de laranja no estado de São Paulo, apesar da importância relativa, observa-se que a produção não se distribui uniformemente entre os citricultores, havendo uma discrepância entre o número de produtores e a quantidade produzida. De acordo com Clemente (2010), a estrutura de produção de laranja no interior paulista caracteriza-se “por muitos produzem pouco” e “poucos produzem muito”. Os produtores de até 100 hectares correspondem a 48% do número total de produtores, mas respondem por apenas 17,5% da produção. No outro extremo, os produtores com mais de 300 hectares correspondem a apenas 17% do número total de produtores, porém respondem com 43,3% da produção paulista. A partir dessas informações acredita-se haver espaço para investigar a existência de uma possível ineficiência na produção de laranjas em São Paulo.

Na verdade, o conceito de eficiência é relativo e difere dos conceitos de eficácia e produtividade. A eficácia está ligada apenas ao que é produzido, sem levar em conta os recursos usados para a produção. A produtividade é denominada pela razão entre o que foi produzido e o que foi gasto para produzir. Já a eficiência compara o que foi produzido, dados os recursos disponíveis, com o que poderia ter sido produzido com os mesmos recursos, de modo que, caso a unidade produtiva esteja muito distante desse parâmetro, ela pode ser considerada ineficiente. Existem duas formas de uma unidade não eficiente tornar-se eficiente: a primeira é reduzindo os insumos, mantendo constante a produção; a segunda é aumentando a produção, mantendo constantes os insumos (Mello et al. 2005).

Em termos aplicados, a análise de eficiência de unidades produtivas tem importância tanto para fins estratégicos (comparação entre unidades produtivas), quanto para o planejamento (avaliação dos resultados do uso de diferentes combinações de fatores) e para a tomada de decisão (como melhorar o desempenho atual, por meio da análise da distância entre a produção atual e potencial).

Nesse contexto, o trabalho tem por objetivo analisar a eficiência técnica de propriedades citrícolas do estado de São Paulo, tendo como base de informações o período compreendido entre os anos de 2009 e 2010, fase em que foram realizadas as entrevistas com citricultores para obter informações a respeito

da produção de laranjas em cada propriedade.

Para alcançar o objetivo, o artigo está estruturado em quatro seções, além desta introdução. Na segunda seção, apresenta-se o referencial teórico do trabalho e na terceira seção o referencial analítico. Na quarta e na última seção mostra os principais resultados obtidos e as conclusões do estudo, respectivamente.

## 2 Referencial Teórico

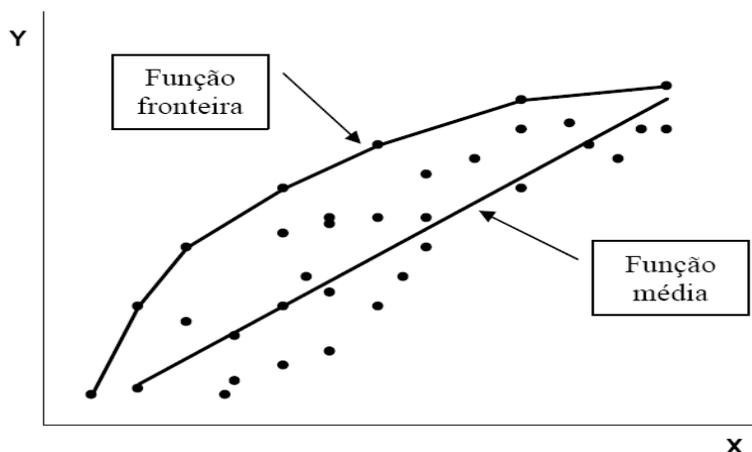
Segundo Carvalho (1984), normalmente, em análises microeconômicas, representa-se a função de produção como sendo uma relação entre  $y$ , a quantidade produzida do bem e um conjunto de parâmetros,  $X_1, X_2, \dots, X_n$ , que identificam as quantidades utilizadas de diversos fatores, respeitando o processo de produção mais eficiente. Em outras palavras, trata-se de “uma relação técnica, que associa a cada dotação de fatores de produção a máxima quantidade de produto obtida a partir da utilização desses fatores” (Barbosa 1985).

Nesse âmbito, a questão da eficiência ganha contornos importantes. Para Sato (1975), a agregação de funções de produção e a sua subsequente estimação econométrica com vistas à geração de uma macro função de produção, sem levar em conta as diferenças na eficiência produtiva, geram resultados viesados. Tais, resultados, ao serem utilizados pelos agentes responsáveis pelo processo produtivo, podem comprometer a alocação eficiente dos recursos, os quais, na maioria das vezes, são escassos e caros.

Em outras palavras, comparando diferentes unidades de produção, pode-se incorrer em erro, caso a análise se baseie apenas na estimação da função de produção média. Isso acontece porque existem diferenças na utilização dos fatores de produção, os quais geram níveis diferentes de eficiência técnica da produção. Assim, para estimar corretamente a função de produção agregada para uma determinada região ou estado, é preciso eliminar as ineficiências existentes em cada unidade produtiva, ou considerá-las adequadamente nas análises pretendidas. Nesse sentido, torna-se necessário estimar uma função de produção de fronteira que caracteriza a melhor tecnologia (*best practice*), a partir da qual pode-se fazer comparações entre unidades de produção em termos de eficiência produtiva e estrutura da tecnologia de produção.

A Figura 1 ilustra a diferença entre uma função de produção média estimada por mínimos quadrados e uma função de produção de fronteira. Percebe-se que na função média, ao minimizar o quadrado dos desvios, existem pontos acima e abaixo da função. Já na função de fronteira, todos os pontos situam-se nela, ou abaixo. Os pontos que se encontram em cima da fronteira referem-se às unidades eficientes. De modo equivalente, os pontos abaixo da fronteira apresentam algum tipo de ineficiência (Fare et al. 1994).

Assim, a existência de ineficiência impossibilita a estimação correta da função, uma vez que contradiz todos os princípios microeconômicos de maximização de lucro. Nesse sentido, é necessário eliminar as ineficiências, no intuito de projetar as unidades ineficientes para a fronteira eficiente. Feito isso, pode-se estimar a função de produção, a qual expressará melhor as relações entre insumos e produto, já desprovida de ineficiência.



Fonte: Gomes (2009)

Figura 1: Representação da função de produção

### 3 Metodologia

#### 3.1 Análise envoltória de dados

Em uma estrutura produtiva, as quantidades máximas de produtos que podem ser obtidas, dados os insumos utilizados, determinam a fronteira de produção (Lins & Meza 2000). A produção de laranja, assim como outras atividades agrícolas, envolve sistemas de produção muito variáveis o que torna mais complexa a tomada de decisão acerca da melhor alocação dos recursos.

A análise envoltória de dados (DEA) é uma técnica não paramétrica que se baseia na programação matemática para analisar a eficiência relativa de unidades produtoras. Na literatura relacionada com modelos DEA, uma unidade produtora é tratada como uma unidade tomadora de decisão (DMU, *decision making unit*), uma vez que desses modelos provém uma medida para avaliar a eficiência relativa de DMUs (Gomes 2009).

Na DEA, o termo análise envoltória deriva do fato de, nessa abordagem, a análise ter como referência as DMUs *outliers*, buscando detectar as DMUs eficientes e construir um plano geométrico de apreciação dessas unidades.

De acordo com Ferreira (2005), a diferença fundamental entre a abordagem DEA e a análise paramétrica, como a fronteira estocástica, está no fato de a primeira ser não paramétrica, estimando uma fronteira determinística, e a segunda ser paramétrica, com base em função estocástica.

Uma significativa limitação ao uso da abordagem paramétrica para mensuração da eficiência decorre do fato de que ela requer que se conheçam, ou se utilizem meios para estimar, estatisticamente, a função de produção considerada. Devido à complexidade das relações de mercado, a estimação da forma funcional apropriada para diferentes DMUs nem sempre é possível. Assim, o enfoque não paramétrico, que utiliza a programação matemática, a exemplo de DEA, parece ser mais apropriado.

Outra considerável vantagem da DEA em relação à estimação paramétrica está na identificação individualizada de cada produtor no quesito eficiência, o que é possível através dos escores de eficiência gerados pela operacionalização do modelo.

Essas características conferem ao método uma potencialidade de explicar, com maior propriedade e com pouca interferência dos avaliadores, as complexidades inerentes às condições reais (Ferreira 2005).

Apesar das vantagens apresentadas, essa metodologia também apresenta desvantagens, dentre as quais se destaca a sensibilidade à presença de outliers e à inclusão ou exclusão de uma ou mais unidades no conjunto de observações, o número de variáveis consideradas na análise, a impossibilidade de se testar estatisticamente os resultados e ainda, a desconsideração da presença de fatores aleatórios e erros de medição, de forma que toda a distância para a fronteira é considerada devido à ineficiência (Nascimento 2013).

Metodologias alternativas ao DEA foram desenvolvidas com objetivo de avançar nos estudos sobre eficiência. Nesse sentido, Aigner et al. (1977) propuseram uma função de produção específica para dados “cross-section”, em que os desvios observados em relação à função de produção poderiam ser devidos à ineficiência produtiva e aos efeitos aleatórios. Essa função é denominada fronteira de produção estocástica (SFA).

O modelo SFA possui vantagens, sendo que as principais são a pouca sensibilidade a problemas de erros de medida, a estimação de intervalos de confiança para os coeficientes de eficiência e a desobrigatoriedade de assumir uma hipótese a respeito de retornos de escala, e desvantagens, como o fato de poder sofrer dos mesmos problemas tradicionais de análise de regressão, de limitações relacionadas à omissão de variáveis, a possível autocorrelação dos erros, a heterocedasticidade e endogeneidade (Nascimento 2013).

Dessa forma, constata-se que não existe consenso na literatura especializada justificando a escolha da DEA ou de fronteiras estocásticas, já que ambas têm vantagens e desvantagens. A escolha do modelo DEA, para este trabalho, deveu-se, principalmente, ao tamanho da base de dados da análise.

No modelo DEA, segundo Charnes et al. (1994), para estimar e analisar a eficiência relativa das DMUs, utiliza-se a definição de ótimo de Pareto, segundo o qual nenhum produto pode ter sua produção aumentada sem que sejam aumentados os seus insumos ou diminuída a produção de outro produto. A eficiência é analisada, relativamente, entre as unidades.

Para incorporar a natureza multiproduto e multi-insumo da produção, Charnes et al. (1994) propõe a técnica DEA para a análise das diferentes unidades, quanto à eficiência relativa.

A função distância<sup>1</sup> é empregada para incorporar a natureza multiproduto e multi-insumo na análise de produtividade e eficiência, sem a necessidade de especificar objetivos comportamentais dos tomadores de decisão.

De acordo com Fare et al. (1994), a forma conveniente de descrever a característica multiproduto da produção é pela tecnologia de produção, definida pelo conjunto  $S$ , representado na equação 1:

$$S = \{(x, y) : x \text{ pode produzir } y\} \quad (1)$$

que é definido pelo conjunto de todos os vetores de insumos e produtos  $(x, y)$ , tal que  $x$  possa produzir  $y$ , em que  $x$  é um vetor  $(k \times 1)$  não negativo de insumos e  $y$ , um vetor  $(m \times 1)$  não negativo de produtos.

<sup>1</sup>Pode ser definida como orientação insumo ou orientação produto. A orientação insumo caracteriza a tecnologia de produção pela minimização proporcional (contração) do vetor insumo, dado um vetor de produto. Já a orientação produto caracteriza a tecnologia de produção pela maximização proporcional do vetor produto, dado um vetor de insumo.

O conjunto de tecnologias de produção pode, de forma equivalente, ser definido pelo conjunto de possibilidades de produção  $P(x)$ , que representa o conjunto de todos os vetores de produtos  $y$ , que pode ser produzido pelo vetor de insumos  $x$ , isto é,

$$Px = \{y : x \text{ pode produzir } y\} \quad (2)$$

A função distância com orientação produto, de acordo com Shephard (1970), pode ser definida pelo conjunto de produtos  $P(x)$ , como

$$d_0(x, y) = \min \left\{ \phi : \frac{y}{\phi} \in P(x) \right\} \quad (3)$$

$$d_0(x, y) = (\max \{ \phi : (\phi y) \in P(x) \})^{-1}$$

em que  $\phi$ , na expressão 3, é um fator mínimo, pelo qual o produto pode ser contraído e, ainda assim, pertencer ao conjunto de possibilidades de produção.

A função distância  $d_0(x, y)$  poderá ter valores menores ou iguais a 1, se o vetor de produto  $y$  for um elemento do conjunto de possibilidade de produção  $P(x)$ ; se for igual a 1,  $(x, y)$  estará sobre a fronteira tecnológica; nesse sentido, a produção será tecnicamente eficiente.

O modelo DEA com orientação-produto e pressuposição de retornos não constantes à escala procura maximizar o aumento proporcional nos níveis de produto, mantendo fixa a quantidade de insumos. De acordo com Charnes et al. (1994), pode ser representado algebricamente por:

$$[d_0(x, y)]^{-1} = \text{MAX}_{\theta, \lambda, S^+, S^-} \phi$$

sujeito a

$$\phi y_i - Y\lambda + S^+ = 0$$

$$-x_i + X\lambda + S^- = 0$$

$$N1' \lambda \leq 1 \quad (4)$$

$$\lambda \geq 0$$

$$S^+ \geq 0$$

$$S^- \geq 0$$

em que  $Y_i$  é um vetor ( $m \times 1$ ) de quantidades de produto da  $i$ -ésima DMU;  $x_i$  é um vetor ( $k \times 1$ ) de quantidades de insumo da  $i$ -ésima DMU;  $Y$  é uma matriz ( $n \times m$ ) de produtos das  $n$  DMUs;  $X$  é uma matriz ( $n \times k$ ) de insumos das  $n$  DMUs;  $\lambda$  é um vetor ( $n \times 1$ ) de pesos;  $N1$  é um vetor ( $n \times 1$ ) de números uns;  $S^+$  é um vetor de folgas relativo aos produtos;  $S^-$  é um vetor de folgas relativos aos insumos; e  $\phi$  é um escalar que tem vetores iguais ou maiores do que 1 e indica o escore de eficiência das DMUs, ou seja, um valor igual a 1 indica eficiência técnica da  $i$ -ésima DMU, em relação às demais, enquanto um valor maior que 1 evidencia a presença de ineficiência técnica relativa. O problema apresentado em 4 é resolvido  $n$  vezes - uma vez para cada DMU, e, como resultado, apresenta os valores de  $\phi$  e  $\lambda$ , sendo  $\phi$  o escore de eficiência da DMU sob análise e  $\lambda$  fornece os *peers* (as DMUs eficientes que servem de referência para a  $i$ -ésima DMU ineficiente).

### 3.2 Modeto Tobit

Para verificar quais variáveis estão associadas à eficiência e à ineficiência das propriedades citrícolas do estado de São Paulo, utiliza-se o modelo econométrico Tobit.

Segundo Ferreira (2005), o modelo Tobit é utilizado nos casos em que a variável dependente está compreendida entre certos valores ou concentrada em pontos iguais a um valor-limite. Conforme Vasconcellos & Alves (2000), isso se dá em decorrência de os fatores ligados à mensuração dos dados não possibilitarem a representação de valores da variável de respostas para toda a sua extensão. Nessas situações, a aplicação do modelo Tobit objetiva contornar o problema da censura valendo-se de técnicas estatísticas que possibilitem fazer inferências para toda a população sem perda de qualidade.

De acordo com Greene (1993), quando se pretende incorporar o erro novamente na equação para se computar a parte estocástica, haverá, em alguns casos, a necessidade de se trabalhar com modelos de variáveis censuradas ou, como mais conhecido, modelo Tobit cujo nome se deve a James Tobin, seu criador.

O modelo Tobit é estimado da seguinte forma:

$$Y_i^* = \beta X_i' + \mu_i \quad (5)$$

onde  $\mu_i$  é iid  $\approx N(0, \sigma^2)$ .

Os valores observados de  $Y_i$  são chamados de  $Y^*$ , ou seja, são condicionadas a:

$$\begin{aligned} Y_i &= Y^* \text{ se } Y_i^* \geq Y_0 \\ Y_i &= Y_0 \text{ se } Y_i^* \leq Y_0 \end{aligned} \quad (6)$$

Conforme colocado por Amemiya (1984), a função de ML para o modelo Tobit é:

$$L = \prod_0 [1 - \phi(X' \alpha)] \prod_1 \sigma^{-1} \phi \left[ \left( \frac{Y_i}{\sigma} - X_i' \alpha \right) \right] \quad (7)$$

Na equação 7, o primeiro membro representa a probabilidade de o evento ser menor que zero e o segundo termo representa a densidade do evento ser observado. Nos dois termos  $\alpha = \frac{\beta}{\sigma}$ , pois é uma forma conveniente de escalar a equação para convergência.

Nesse ponto, procura-se concentrar em informações referentes às características das propriedades (tamanho e número de funcionários) e característica do produtor (idade, escolaridade e tempo como produtor). Espera-se que essas variáveis impactem positivamente no fato das propriedades serem eficientes. Assim a seguinte equação foi estimada, com base nos dados primários obtidos com uma amostra de citricultores do estado de São Paulo:

$$Y_i = \beta_1 + \beta_2 I_i + \beta_3 E_i + \beta_4 T e_i + \beta_5 T_i + \beta_6 N f_i + \epsilon_i \quad (8)$$

em que:

$Y_i$  = escores de eficiências obtidos por meio da Análise envoltória de dados. Logo, cada DMU tem um coeficiente positivo de eficiência, limitada ao intervalo de 0 a 1;

$I_i$  = idade do produtor (em anos);  
 $E_i$  = escolaridade do produtor (em anos);  
 $Te_i$  = tempo como produtor rural (em anos);  
 $T_i$  = tamanho da propriedade (em hectares);  
 $Nf_i$  = número de funcionário nas propriedades;  
 $\epsilon_i$  = termo de erro.

A estimação da equação 8 permite fazer inferências para toda a população sem perda de qualidade.

Espera-se que todas essas variáveis impactem positivamente na eficiência da propriedade. As variáveis idade e tempo como produtor rural evidenciam a experiência do produtor no ramo da citricultura. A variável “escolaridade” indica o nível de conhecimento técnico do citricultor, e a variável “tamanho da propriedade” transfigura na produção de laranjas concretizada.

### 3.3 Função de produção

Para analisar a produção, utiliza-se a função de produção tipo *Cobb-Douglas*, definida da seguinte forma:

$$Y = AX_1^{\beta_1} X_2^{\beta_2} X_3^{\beta_3} \quad (9)$$

em que  $Y$  é a variável dependente relacionada à produção;  $A$  indica o nível tecnológico;  $X_i$  refere-se às variáveis independentes, representando os fatores de produção terra, trabalho e capital; e os  $\beta_i$  são os parâmetros a serem estimados.

Aplicando-se logaritmo em ambos os lados da equação, define-se a forma funcional da função de produção média, descrita pelo seguinte modelo de regressão:

$$\text{Log } Y_i = \text{Log } A + \beta_1 \text{Log } X_{1i} + \beta_2 \text{Log } X_{2i} + \beta_3 \text{Log } X_{3i} + \mu_i \quad (10)$$

em que os  $\beta$  s referem-se aos parâmetros a serem estimados; e  $\mu_i$  representa o termo de erro aleatório, que se pressupõe ter média 0 e variância constante. Ao expressar a função dessa forma, os coeficientes  $\beta_i$  passam a ser as elasticidades de produção de cada fator.

Nessa função de produção, estimada com os dados originais, existem ineficiências relativas entre os produtores da amostra. Com isso, após calcular o escore de eficiência para cada produtor, os valores projetados das variáveis são utilizados para estimar uma nova função de produção agregada, já descon sideradas as ineficiências na alocação de recursos no processo produtivo. Essa nova função, denominada de função de produção de fronteira, foi estimada de forma semelhante à função média, diferindo apenas nas variáveis utilizadas, isto é,

$$\text{Log } Y_i^P = \text{Log } A + \beta_1 X_{1i}^P + \beta_2 X_{2i}^P + \beta_3 X_{3i}^P + \mu_i \quad (11)$$

em que:

$$\begin{aligned} Y_i^P &= \phi Y_i + S_i^+ \\ X_{1i}^P &= X_{1i} - S_{1i}^- \\ X_{2i}^P &= X_{2i} - S_{2i}^- \\ X_{3i}^P &= X_{3i} - S_{3i}^- \end{aligned} \quad (12)$$

em que  $\phi_i$  é o escore e eficiência estimado por meio do modelo de envoltória apresentado em 4 para o  $i$ -ésimo produtor e  $S_{ji}$  refere-se à folga na  $j$ -ésima variável estimada para cada produtor.

### 3.4 Área de Estudo

Com a finalidade de investigar o nível de produção de laranja dos citricultores bem como a quantidade de insumos utilizados, foram realizadas entrevistas com produtores de laranja do estado de São Paulo previamente selecionados para aplicação de um questionário estruturado, no período de dezembro de 2009 a fevereiro de 2010. Foram entrevistados citricultores de onze cidades do interior do estado de São Paulo, região que concentra a maior produção de laranja do país.

O estado de São Paulo possuía, em 2006, aproximadamente 6300 citricultores ativos, concentrando 81% de toda a produção nacional de laranja.

Neves (2005) evidencia que as regiões citrícolas em São Paulo dividem-se em norte e nordeste, centro, sul e sul novo (Figura 2). A primeira compreende a região de Bebedouro e Barretos, São José do Rio Preto e Votuporanga e a região de Catanduva, responsável por 45% da produção do estado. Na segunda, participam a região de Araraquara e Matão, Itápolis e Taquaritinga, representando 30% da produção. No sul, fazem parte a região de Limeira, Avaré/Botucatu e Itapetininga. Já o sul novo contempla de Bauru a Itapetininga. Essas duas últimas regiões concentram 25% da produção (Tavares 2006).



Fonte: Alicitros (2009)

**Figura 2:** Principais regiões produtoras do estado de São Paulo

Ao se estruturar a amostra, o objetivo foi selecionar citricultores das regiões onde mais se concentra a produção de laranja do estado de São Paulo. O cálculo do tamanho da amostra<sup>2</sup>, com um intervalo de confiança e tolerância do erro amostral de 10%, resultou em 67 questionários.

Para a estratificação da amostra, foram utilizadas a participação das regiões na produção de laranja do estado de São Paulo e, em seguida, a produção das principais regiões em 2003. Assim, selecionou-se 30 questionários para a

<sup>2</sup>Cálculo do tamanho da amostra:  $n = \frac{Z^2 \cdot p \cdot q \cdot N}{[d^2 \cdot (N-1) + Z^2 \cdot p \cdot q]}$  (Greene 1993).

região norte e nordeste, 20 questionários para a região central e 17 questionários para a região sul/sul novo. Como as regiões de Barretos, Catanduva e São José do Rio Preto se localizam no norte e nordeste, foram aplicados 17 questionários para a região de Barretos, 6 para a região de Catanduva e 7 para a região de São José do Rio Preto. Para o centro, os 20 questionários foram aplicados na região de Araraquara e para o sul/sul novo foram aplicados 10 questionários na região de Limeira, 4 na região de Jaú e 3 na região de Bauru.

### 3.5 Fonte de dados

Parte importante da construção da base de dados utilizada nesta pesquisa baseou-se em uma pesquisa exploratória com 67 produtores de laranja localizados em onze municípios paulistas, entre os meses de janeiro e março de 2010. Nessa ocasião, foram levantadas as informações sobre as variáveis utilizadas, como idade, escolaridade e tamanho da propriedade, dentre outras.

Os dados relativos ao custo médio de capital das propriedades do interior do estado de São Paulo e o preço médio da laranja vendida para indústria foram coletados junto ao Instituto de Economia Agrícola (IEA).

## 4 Resultados e Discussão

### 4.1 Características descritivas da amostra utilizada

Para realização da pesquisa, foram aplicados questionários estruturados a 67 produtores de laranja do estado de São Paulo. Ao se estruturar a amostra, foram selecionados produtores dos principais municípios paulistas. O objetivo foi selecionar citricultores das regiões onde mais se concentra a produção de laranja do estado de São Paulo.

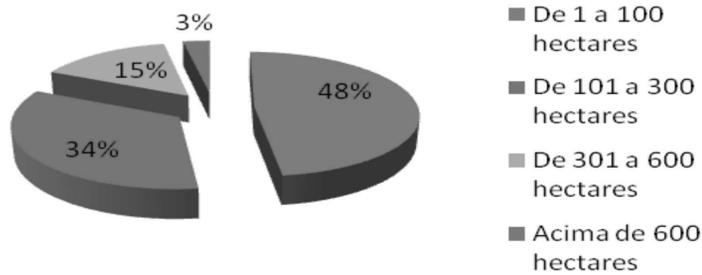
Dentre as principais características observadas, notou-se que, no total, 34% dos entrevistados têm idade entre 23 e 50 anos e 66% estão com idade acima de 50 anos, o que indica a predominância de produtores com idade mais avançada, o que corrobora a ideia do relativo envelhecimento dos produtores brasileiros, já identificado em outros trabalhos.

Verificou-se, também, razoável nível de escolaridade, já que 33% dos entrevistados possuem até 5 anos de estudo e 51% possuem mais de 10 anos de estudo. Em relação ao “tempo como produtor rural”, os resultados mostram que 26% dos entrevistados têm até 20 anos na atividade, 24% dos produtores têm entre 20 e 30 anos e 50% produzem a mais de 30 anos, evidenciando a predominância de citricultores com larga experiência na produção da laranja.

A partir da análise do tamanho da propriedade dos citricultores (Figura 3), destaca-se a presença de “pequenas” e “médias” propriedades (até 100 hectares), com 48%. Para 81% dos produtores, a principal fonte de renda é a agricultura e 55% obtêm uma renda bruta anual com a laranja acima de R\$ 100.000 (Figura 4).

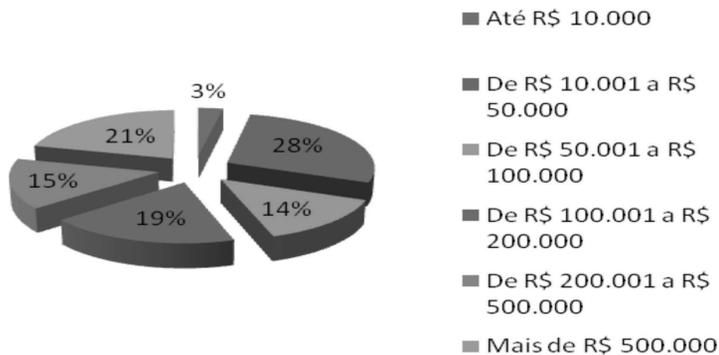
Quanto ao perfil da mão de obra, os resultados mostraram que 84% das propriedades possuem mão de obra contratada e apenas 16% contam com mão de obra exclusivamente familiar. Embora a utilização de mão de obra familiar seja uma característica de pequenas propriedades, verifica-se, também, a presença de mão de obra permanente e remunerada.

Com relação ao número de empregados permanentes, verifica-se que a média de trabalhadores contratados por propriedade é de, aproximadamente,



Fonte: Clemente (2010)

**Figura 3:** Tamanho da propriedade



Fonte: Clemente (2010)

**Figura 4:** Renda bruta anual com a laranja

seis. Porém, a desigualdade da mão de obra entre as fazendas é muito grande, pois há propriedades com até 75 funcionários e propriedades com nenhuma. Isso também é evidenciado pelo alto valor do desvio-padrão (11,75).

Em síntese, os resultados do levantamento primário apontam diferenças significativas nas características produtivas das propriedades citrícolas do estado de São Paulo, com destaque para a escolaridade, a renda bruta e o tempo como produtor rural. Essas divergências podem fomentar algum grau de ineficiência na produção de laranja da região.

#### 4.2 Análise de Eficiência Técnica

De maneira geral, uma análise de eficiência técnica compara o que foi produzido, dados os recursos disponíveis, com o que poderia ter sido produzido com os mesmos recursos. No caso em estudo, as variáveis empregadas no modelo de eficiência para a amostra como um todo, juntamente com suas estatísticas descritivas são apresentadas na Tabela 1. Essas variáveis refletem características da propriedade (produção, tamanho, funcionários e custo de capital) e do produtor (idade, escolaridade e tempo como produtor). Percebe-se uma relativa diferença de magnitude entre as unidades que compõem a

amostra, destacando, principalmente, o alto desvio-padrão resultante da relativa dispersão dos dados em torno da média, que enfraquece as inferências de tendência central.

**Tabela 1:** Estatísticas descritivas das variáveis empregadas no modelo de eficiência dos produtores, 2010

Variável	Média	Desvio-padrão	Máximo	Mínimo
Produção de laranja (cx 40,8kg)	34920,59	39444,87	136263,00	738,37
Tamanho da propriedade (em ha)	190,66	287,95	2100,00	12
Número de funcionários	6,13	11,75	75	0
Custo de capital (em R\$)	714063,98	1049592,85	7578522,00	46695,12
Idade do produtor (em anos)	55,21	13,11	79	23
Escolaridade do produtor (em anos)	9,93	5,43	17	1
Tempo como produtor (em anos)	31,52	14,29	63	5

Na Tabela 2, é apresentado o sumário estatístico do cálculo da eficiência técnica das unidades que compõe a amostra. Pela média dos escores de eficiência técnica, é possível visualizar o grau de eficiência das propriedades. Já os escores individualizados permitem apontamentos mais específicos a cada unidade produtiva, indicando ineficiência em recursos, bem como apontando DMUs que lhe sirvam de modelo. Essa verificação é importante para analisar em qual situação encontram-se cada propriedade em detrimento do desempenho do conjunto.

**Tabela 2:** Escores de eficiência técnica das propriedades citrícolas

Variável	Unidades eficientes	Média	Desvio-padrão	Máximo	Mínimo
Eficiência técnica	13,43%	0,79	0,21	1,00	0,25

Os resultados demonstram que as propriedades citrícolas de São Paulo apresentam expressivo grau de ineficiência técnica.

A média da eficiência técnica foi de 0,79, o que sugere a possibilidade de aumento de produção, considerando a mesma proporção de insumos atualmente utilizada, tomando como referência a orientação produto do modelo. Ao analisar as regiões produtoras do estado de São Paulo, evidencia-se que 67% das propriedades eficientes encontram-se no norte e nordeste do estado. Isso ocorre devido ao fato dessas regiões serem as mais antigas na produção de laranja do país, o que possibilitou aos produtores maior conhecimento quanto à melhor combinação de insumos. Com relação às propriedades mais ineficientes, 71,4% estão localizadas na região sul e sudeste do estado.

Para comparar o grau de ineficiência da unidade produtiva tomando por base o escore da média de eficiência técnica, foi construído o indicador definido por Ferrier & Porter (1991), que segue:

$$\left( \frac{1}{score} - 1 \right) \times 100 \quad (13)$$

Assim, é possível dizer que a diferença média de eficiência técnica nas propriedades de laranja requer uma utilização de recursos 26,6% maior que as propriedades que estão operando sobre a curva de produção.

Com o objetivo de verificar quais os determinantes da eficiência das propriedades citrícolas do estado de São Paulo, utilizou-se o modelo econométrico Tobit. Os resultados do modelo podem ser observados na Tabela 3.

**Tabela 3:** Fatores associados à eficiência técnica nas propriedades citrícolas do estado de São Paulo — método Tobit

Variável	Coefficiente	Erro-padrão	Valor-p
Tamanho da propriedade (em ha)	0,00021	0,00012	0,096
Número de funcionários	-0,0057	0,00292	0,054
Idade do produtor (em anos)	0,0017	0,00183	0,365
Escolaridade do produtor (em anos)	0,0283	0,00336	0
Tempo como produtor rural (em anos)	0,0039	0,00168	0,02
Constante	0,2866	0,09289	0,003
$\chi^2 = 55,48$			

Os coeficientes estimados por **meio do modelo** apresentaram ajustamento satisfatório, conforme verificado pelo nível de significância do teste  $\chi^2$ , que foi significativo em 1% de probabilidade.

Dentre os fatores que tiveram influência na eficiência técnica, ressalta-se a variável “escolaridade do produtor”, que mais impacta na eficiência das propriedades citrícolas. Isso mostra que a capacitação é o principal meio para obter eficiência técnica na produção de laranja.

A variável “tempo como produtor rural” também influencia diretamente a eficiência nas propriedades, indicando que a experiência com a produção é uma variável importante para determinar a eficiência dos citricultores.

A variável “idade”, embora importante, não foi significativa. Isso pode ter ocorrido devido à forte relação desta com a variável “tempo como produtor rural”. A variável “número de funcionários permanentes” mostrou-se significativa com sinal contrário ao esperado. Já a variável “tamanho da propriedade” foi significativa e positiva, porém com um impacto muito pequeno na determinação da eficiência citrícola.

Com isso, é possível verificar que, para as propriedades produtoras de laranja do estado de São Paulo, além da experiência adquirida como produtor rural, evidencia-se a importância da educação formal que objetiva aprimorar as técnicas dos produtores, promovendo, assim, o aumento da eficiência na produção do fruto.

Embora o uso de modelos Tobit como “segundo estágio”, para explicar índices de eficiência provenientes da estimação de fronteiras na DEA, tenha ganhado popularidade nas décadas de 1990 e de 2000, mais recentemente McDonald (2009) demonstrou que seu uso pode ser inapropriado, e que, em tais aplicações, o estimador de Máxima Verossimilhança (ML) é geralmente inconsistente, ao contrário do estimador de Mínimos Quadrados Ordinários (MQO). De forma a constatar ou não essa evidência, estimou-se novamente a regressão, utilizando o método de MQO, conforme Tabela 4.

Observa-se que, embora os valores dos parâmetros sejam bem semelhantes e com os mesmos sinais apresentados na Tabela 3, o erro-padrão e, consequentemente, a significância estatística se mostram bem distintas. Enquanto no primeiro modelo (ML) as variáveis “tamanho da propriedade” e “tempo como produtor rural” foram significativas a 10%, no segundo modelo (MQO) essas passaram a ser não significativas. O contrário ocorre com a variável “idade do produtor”, que passa a ser significativa a 15% no segundo modelo.

**Tabela 4:** Fatores associados à eficiência técnica nas propriedades citrícolas do estado de São Paulo — método Mínimos Quadrados Ordinários (MQO)

Variável	Coefficiente	Erro-padrão	Valor-p
Tamanho da propriedade (em ha)	0,00028	0,00026	0,28
Número de funcionários	-0,0101	0,00617	0,106
Idade do produtor (em anos)	0,0063	0,00388	0,107
Escolaridade do produtor (em anos)	0,0145	0,0071	0,045
Tempo como produtor rural (em anos)	0,0048	0,00354	0,178
Constante	1,391	0,19623	0
$R^2 = 22,85$			

Para verificar possível inconsistência e tendenciosidade do MQO, pensou-se em estimar o modelo anterior retirando as observações cujo valor é zero ( $y = 0$ ) e compará-lo com os parâmetros verdadeiros. Esse resultado encontra-se na Tabela 5. Pode-se observar que as estimativas são diferentes das apresentadas na Tabela 4, o que mais uma vez justifica o uso de um estimador de máxima verossimilhança, que é consistente e usa todas as observações.

**Tabela 5:** Fatores associados à eficiência técnica nas propriedades citrícolas do estado de São Paulo — método Mínimos Quadrados Ordinários (MQO)

Variável	Coefficiente	Erro-padrão	Valor-p
Tamanho da propriedade (em ha)	0,00024	0,00027	0,373
Número de funcionários	-0,0085	0,00631	0,178
Idade do produtor (em anos)	0,0059	0,00393	0,138
Escolaridade do produtor (em anos)	0,0177	0,00744	0,021
Tempo como produtor rural (em anos)	0,0051	0,00364	0,169
Constante	1,426	0,1993	0
$R^2 = 24,66$			

### 4.3 Função de Produção

No intuito de eliminar as ineficiências existentes em cada unidade produtiva, foram estimadas funções de produção média e de fronteira do tipo Cobb-Douglas. A forma funcional utilizada foi escolhida por permitir identificar, diretamente, a elasticidade de produção de um fator, a qual também indica a importância daquele fator no processo produtivo. A função estimada utilizando-se os dados originais foi denominada “função média”, cujos resultados encontram-se na segunda coluna da Tabela 6.

Conhecendo-se a medida de eficiência e as folgas dos insumos para cada propriedade, elimina-se dos dados originais os excessos de uso dos fatores (ineficiência), obtendo-se novos valores para as variáveis. Com isso, estimou-se a “função de produção de fronteira”, cujos resultados também se encontram na Tabela 6.

Analisando os resultados das duas funções, nota-se que existem diferenças significativas nos parâmetros da função média e da função fronteira. A variável “terra” mostra-se mais importante na função média, seguida da variável “trabalho”. Já para função de produção de fronteira, a variável “trabalho”

**Tabela 6:** Coeficientes estimados das elasticidades de produção e  $R^2$  corrigido para as funções de produção *Cobb-Douglas*

Coeficientes	Modelo estimado		Variação
	Função Média	Função Fronteira	
Intercepto	55,2430***	3,0920**	-
$\beta_1$ (terra)	6,0120***	0,5510*	- 91%
$\beta_2$ (trabalho)	0,5740*	0,6530*	14%
$\beta_3$ (capital)	- 5,7440***	0,1240 <sup>NS</sup>	-102%
$R^2$ corrigido	47%	59%	-

evidencia-se mais importante, seguida da variável “terra”. Ambas tiveram no capital um fator não significativo.

Embora o trabalho e a terra tenham suas importâncias relativas no processo produtivo, observa-se que existem grandes diferenças na magnitude dos coeficientes entre as duas funções. O coeficiente do fator terra varia em 91% quando se compara as funções de produção. Já o coeficiente do fator trabalho varia 14% entre as funções. De certa forma, essas diferenças indicam o tamanho do erro que se pode incorrer na formulação de políticas para o setor caso não utilize uma função de produção corrigida pelas ineficiências. Também, pela análise das diferenças entre os parâmetros das duas funções de produção tem-se uma ideia do impacto da ineficiência na produção.

Gomes & Baptista (2009) realizaram o cálculo da função média e de fronteira para o setor agropecuário de Minas Gerais. Como principal resultado destaca-se as diferenças significativas nos parâmetros da função média e da função fronteira. De acordo com os autores, observar-se que a função de produção estimada a partir dos dados originais contém parcela significativa de ineficiência produtiva, invalidando a análise. Para saber o real comportamento da produção agrícola é necessário estimar funções de produção de fronteira, que considera os dados desprovidos de ineficiência.

Em relação à elasticidade de escala, percebe-se que o procedimento utilizado para estimar a função de produção possui uma característica importante, que é o fato de impor restrições quanto à homogeneidade da função de produção. Em outras palavras, ao utilizar o modelo de envoltória de dados com pressuposição de retornos não constantes à escala, impõe-se a concavidade da função de produção de fronteira, corrigindo a ineficiência produtiva e impondo restrições na estrutura da tecnologia de produção.

Conforme mostra a literatura, Alves (2002) evidencia que a utilização da função do tipo Cobb-Douglas para produtos agrícolas apresenta uma vantagem do ponto de vista da agregação, devido ao fácil ajustamento. Entretanto, a utilização da função Cobb-Douglas é muito restritiva, pois supõe elasticidade de substituição constante e unitária e só é justificada em casos de multilinearidade elevada. Para tentar corrigir esse problema, poderia utilizar uma forma mais flexível, a forma translog, que também se torna linear por logaritmos, e é comumente utilizada em aplicações agrícolas (Albuquerque 1987).

## 5 Conclusão

Neste estudo, objetivou-se analisar a eficiência das propriedades citrícolas do estado de São Paulo, no período de 2009 a 2010. Com isso, aplicou-se a abordagem não paramétrica de análise envoltória de dados para calcular os níveis de eficiência técnica. Para conhecer os determinantes da eficiência técnica dos citricultores, utilizou-se o modelo econométrico Tobit e, posteriormente, para evidenciar as diferenças entre as funções de produção média e de fronteira e auxiliar no processo de tomada de decisão no setor, ajustaram-se funções de produção agregada na forma funcional tipo *Cobb-Douglas*.

Com base nos resultados, confirma-se a hipótese de que as propriedades citrícolas atuam de forma ineficiente, ou seja, não alocam de forma ótima seus recursos para a produção. Para elevar a eficiência na produção, faz-se necessário aumentar a educação formal dos produtores, de forma a possibilitar maior conhecimento técnico para a produção. As funções de produção média e de fronteira mostraram-se diferentes quanto à magnitude dos seus parâmetros, indicando que esse instrumental de análise é muito importante e sua estimação e interpretação, de forma correta, pode facilitar a tomada de decisão e melhorar significativamente a alocação dos recursos.

Apesar da divergência na literatura acerca do melhor método para o “segundo estágio”, constata-se que o uso do método de máxima verossimilhança é o mais adequado para esses dados, haja vista que a variável dependente se apresenta com uma concentração de observações no limite inferior.

Assim, conclui-se que é importante que os agentes envolvidos com o setor produtivo da laranja tenham consciência da existência de ineficiências. Em muitos casos, a redução dessas ineficiências poderia melhorar a alocação dos recursos em todo o setor, ou seja, os insumos utilizados em excesso em algumas propriedades podem ser realocados, aumentando, com isso, o volume total da produção.

Uma sugestão para trabalhos futuros seria o cálculo da função de produção *translog* e comparação com a função de produção *Cobb-Douglas*.

## Referências Bibliográficas

Agriannual (2010), *Anuário Estatístico da Agricultura Brasileira*, Argos, São Paulo.

Aigner, D., Lovell, K. & Schmidt, P. (1977), ‘Formulation and estimation of stochastic frontier production function models’, *Journal of Econometrics* 6, 21–37.

Albuquerque, M. (1987), ‘Uma análise *translog* sobre mudança tecnológica e efeitos de escala: um caso de modernização ineficiente’, *Pesquisa e Planejamento Econômico* 17, 191–220.

Alicitros, A. d. C. d. R. d. L. (2009), ‘Histórico da laranja’.

URL: <http://www.alicitros.com.br/index2/principal.php?p=laranja>

Alves, L. (2002), ‘Produção agrícola agregada do estado do paran  em 1995’, *Anhanguera Goi nia* 3, 103–121.

Amemiya, T. (1984), ‘Tobit models, a survey’, *Journal of Econometrics* 24, 3–61.

- Barbosa, F. (1985), *Microeconomia: teoria, modelos econométricos e aplicações à economia brasileira*, IPEA/INPES, Rio de Janeiro.
- Carvalho, L. (1984), *Teoria da firma a produção e a firma: manual de introdução à economia*, Saraiva, São Paulo.
- Charnes, A., Cooper, W., Lewin, A. & Seiford, L. (1994), *Data envelopment analysis: theory, methodology and application*, Kluwer Academic, Dordrecht.
- Clemente, F. (2010), 'Viabilidade de implantação de contratos futuros de suco de laranja concentrado congelado no Brasil', Viçosa, MG: UFV, IMPRENSA Universitária.
- Fare, R., Grosskopf, S. & Lovell, C. (1994), *Productions frontiers*, Cambridge University, Cambridge.
- Ferreira, M. (2005), 'Eficiência técnica e de escala de cooperativas e sociedade de capital na indústria de laticínios do Brasil', Viçosa, MG: UFV, IMPRENSA Universitária.
- Ferrier, G. & Porter, P. (1991), 'The productive efficiency of US milk processing co-operatives', *Journal of Agricultural Economics* pp. 161–173.
- Gomes, A. P. (2009), *Introdução à análise envoltória de dados: teoria, modelos e aplicações*, Ed. UFV, Viçosa.
- Gomes, A. P. & Baptista, A. (2009), 'Função de produção de fronteira e tomada de decisão na agropecuária', Palestra.
- Greene, W. H. (1993), *Econometric Analysis*, 2 ed. edn, New York: Macmillan. 791p.
- Lins, M. & Meza, L. (2000), *Análise envoltória de dados a perspectiva de integração no ambiente de apoio à decisão*, COPPE/UFRJ, Rio de Janeiro.
- McDonald, J. (2009), 'Using least squares and tobit in second stage DEA efficiency analyses', *European Journal of Operational Research* pp. 792–798.
- Mello, J., Meza, L., Gomes, E. & Neto, L. (2005), 'Curso de análise de envoltória de dados', XXXVII Simpósio Brasileiro de Pesquisa Operacional.
- Nascimento, A. (2013), 'Regulação no setor de distribuição de energia elétrica no Brasil: análise de metodologias alternativas para a definição dos custos operacionais eficientes', Universidade Federal de Viçosa.
- Neves, M. F. (2005), *Estratégias para a laranja no Brasil*, Atlas S.A., São Paulo.
- Sato, K. (1975), *Production functions and aggregation*, North-Holland publishing company, Amsterdam.
- Tavares, M. (2006), 'O mercado futuro de suco de laranja concentrado e congelado: um enfoque analítico', Porto Alegre, RS. UFRGS.
- Vasconcellos, M. & Alves, D. (2000), *Manual de econometria: nível intermediário*, Atlas S.A., São Paulo.



# REAVALIANDO A RELAÇÃO ENTRE PRODUTIVIDADE E EXPORTAÇÕES ATRAVÉS DE MODELO DE CORREÇÃO DE ERROS NÃO LINEAR: EVIDÊNCIAS PARA ESTADOS UNIDOS, CANADÁ, JAPÃO E ALEMANHA

JOÃO PAULO MARTIN FALEIROS \*  
DENISARD CNEIO DE OLIVEIRA ALVES †

## Resumo

O presente artigo avalia a relação entre exportações e produtividade total dos fatores (PFT) através do modelo de correção de erros não linear MR-STVEC (Multiple Regime Smooth Transition VEC) aplicado a quatro países desenvolvidos (Estados Unidos, Canadá, Japão e Alemanha). O MR-STVEC, assim como os demais da família de modelos de transição suave, exige uma estratégia prévia de identificação. Tomando como base a estratégia de Lundbergh *et al.* (2003), são aplicados dois diferentes testes de razão de Máxima Verossimilhança. Os resultados indicam que, para todos os casos, as exportações possuem mecanismo de reverter possíveis choques negativos de produtividade. Adicionalmente, para o Canadá e Alemanha, quando há choques positivos da produtividade, as exportações também agem, mas de modo a restringi-los.

**Palavras-chave:** Exportações; Produtividade; Modelos não lineares.

## Abstract

The aim of this paper is to evaluate the relationship between exports and productivity using Multiple Regimes Smooth Transition Vector Error-Correction Model (MR-STVEC) for a sample of four developed countries (United States, Canada, Japan and Germany). MR-STVEC models require a previous strategy of identification. Based on the strategy used by Lundbergh *et al.* (2003), we apply two different likelihood ratio tests. The results indicate that exports may reverse negative productivity shocks. In particular, for Canada and Germany, exports are able to act as a restriction to an increase in productivity when there are positive shocks.

**Keywords:** Exports; Productivity; Nonlinear models.

**JEL classification:** C22, C51, F43.

**DOI:** <http://dx.doi.org/10.1590/1413-8050/ea120727>

\* Banco Nacional de Desenvolvimento Econômico e Social (BNDES). E-mail: [jpfaleiros@hotmail.com](mailto:jpfaleiros@hotmail.com)

† Faculdade de Economia e Administração da USP (FEA-USP). E-mail: [dcoalves@usp.br](mailto:dcoalves@usp.br)

## 1 Introdução

Em geral, a literatura aponta que, para alguns países, essas séries cointegram e a produtividade é destaque na literatura econômica (Krueger 1980, Bruton 1989, Edwards 1993, 1998, Bernard & Jensen 1999, Vamvakidis 2002). Entretanto, o que ainda é motivo de fortes controvérsias é a direção de causalidade pela qual essa relação vem a ocorrer.

Por um lado está a corrente econômica tradicional que advoga que a expansão das exportações orienta a produtividade de um país (Balassa 1978, Feder 1982, Esfahani 1991, Marin 1992, Rodrik 1995, De Loecker 2007). De maneira geral, a expansão das exportações pode ser um canal pelo qual a nação pode auferir ganhos de competitividade e lucratividade. Esse fato viria a ocorrer por meio da melhoria na alocação de recursos do país, de acordo com sua vantagem comparativa, tal como pela exploração de economias de escala. Outro benefício ligado à expansão das exportações é a difusão de conhecimento através de “spillovers” tecnológicos gerado pela forte competição no mercado mundial. Quando as firmas têm contato com o mercado internacional, passam a adquirir conhecimento e expertise técnica que permitem a melhoria de seu nível de eficiência e produtividade, ou seja, um mecanismo de learning-by-exporting.

De outro lado, estão aqueles que questionam esse mecanismo de propagação. Segundo eles, não existe necessariamente dinâmica de crescimento que se processa das exportações para a produtividade. Pelo contrário, a expansão das exportações seria reposta ao aumento da produtividade doméstica (Krugman 1984, Bhagwati 1998, Grossman & Helpman 1991, Edwards 1993, Coe & Helpman 1995). Recentemente alguns estudos no âmbito da firma vêm contestando a dinâmica de learning-by-exporting como fonte de impacto sobre o crescimento, trazendo evidências de que somente as firmas mais produtivas se lançam às exportações, influenciando diretamente sua performance (Aw et al. 1997, Clerides et al. 1998, Bernard & Jensen 1999, Bernard et al. 2007).

Para avaliar o sentido de causalidade entre exportações e produtividade, ou mesmo outra medida de desempenho econômico, boa parte da literatura empírica faz uso da abordagem de séries de tempo, principalmente dos modelos de correção de erros (ECM), associado ao teste de cointegração de Johansen (1988, 1991), metodologia presentes nos principais trabalhos empíricos sobre o tema (Marin 1992, Henriques & Sadorsky 1996, Xu 1996, Yamada 1988, Awokuse 2007). Em geral, a literatura aponta que, para alguns países, essas séries cointegram e a produtividade é a variável responsável pelo ajuste de curto prazo (Greenaway & Sapsford 1994, Giles & Williams 2000a,b). Portanto, choques positivos nas exportações irão se traduzir em crescimento da produtividade, corroborando a visão tradicional.

Em contrapartida, esse resultado traz consigo outro aspecto associado ao padrão de causalidade. De acordo com o mecanismo padrão de correção de erros (ECM), quando há incidência de choques positivos de produtividade, este é dissipado, dado que as exportações não ajustam. Em especial para nações com baixa participação do setor exportador na economia, este não parece ser comportamento crível. Nesse caso, espera-se o contrário, ou seja, que as exportações passem a ser endógenas, de modo que choques de produtividade tenham reflexos sobre o setor exportador. Conseqüentemente, a variável responsável pelo ajustamento de curto prazo alternaria de acordo com o sinal dos choques de produtividade, indicando um mecanismo não linear de correção

de erros.

Este artigo traz evidência de que, primeiramente, existe relação de longo prazo entre a produtividade total dos fatores e receitas de exportações, avaliadas em termos de sua participação no valor total exportado pelo mundo, para quatro países de alta renda per capita e níveis diferentes de participação das exportações no PIB: Estados Unidos, Canadá, Japão e Alemanha. O resultado mais importante é que a dinâmica de ajustamento de curto prazo, dada pelo vetor de correção de erros, é não linear e especificada de acordo com o modelo MR-STVEC (Multiple Regime Smooth Transition VEC), que justamente permite alternância de regimes em função dos choques de produtividade.

Além de corroborar a visão tradicional, de que choques nas exportações promovem aumento da produtividade, será também evidenciado que nem sempre essa última variável é fracamente endógena. Considerando o caso japonês e alemão, choques positivos de produtividade promovem aumento da razão entre a participação das receitas de exportações e o valor total exportado pelo mundo. Já no caso de choques negativos, todos os quatro modelos MR-STVEC estimados indicam que a produtividade continua sendo a variável responsável pelo ajuste, implicando movimento de reversão desta. Por fim, mesmo levando em conta a especificação não linear, os modelos para Canadá e Alemanha trouxeram evidências de dinâmica de restrição da produtividade em resposta a choques positivos.

Vale destacar que a estimação dos modelos MR-STVEC, assim como toda a família desses modelos, exige estratégia de especificação, que neste artigo irá se basear no trabalho de Lundbergh et al. (2003). Como aqui o tratamento é bivariado, serão aplicados dois testes de razão de verossimilhança (LR).

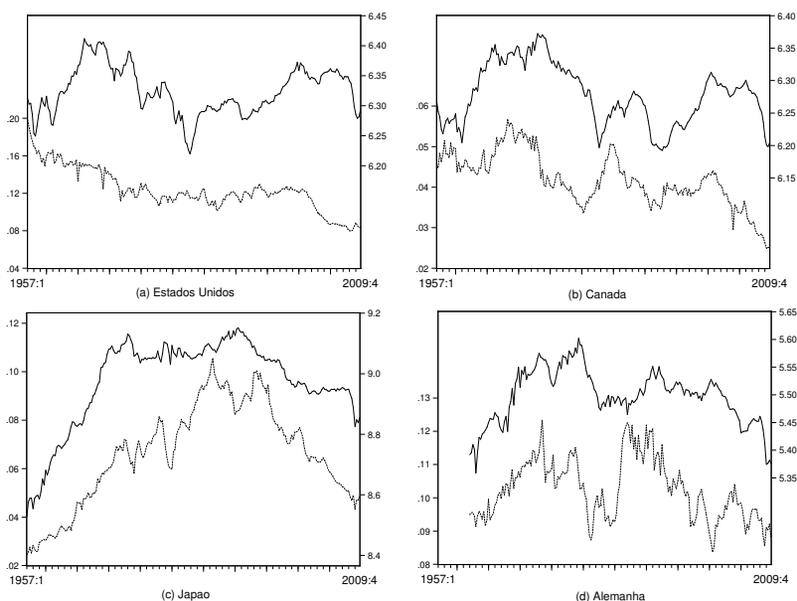
O artigo é organizado em mais cinco seções. A seção 2 trará mais considerações a respeito da relação entre produtividade total dos fatores e exportações. A seção 3 apresenta o modelo MR-STVEC, seguida da seção 4 que abordará a estratégia de especificação. Por sua vez, a seção 5 discutirá o modo como as séries foram construídas e também os resultados obtidos. Finalmente, a seção 6 conclui este trabalho.

## 2 Considerações sobre a relação entre produtividade e exportações

Como destacado, a literatura econômica, de modo geral, não possui muitas restrições de ordem teórica para relacionar exportações à produtividade de um país. Tal como salienta Kunst & Marin (1989) e Marin (1992), até mesmo por meio de breve exame empírico, é possível verificar trajetória conjunta entre exportações e produtividade, incluindo nessa análise, não só países em desenvolvimento, mas aqueles com elevado nível de renda *per capita*. De fato, ao realizar inspeção casual entre as séries da produtividade total dos fatores ( $PTF_t$ ) de quatro países nessa faixa de renda, no caso, Estados Unidos, Canadá, Alemanha e Japão, e a participação das receitas de exportações de bens no valor total exportado mundialmente ( $EX_t$ ), nota-se que as essas duas séries coevoluem ao longo dos últimos 60 anos, como aponta a Figura 1.

A produtividade total dos fatores foi calculada como resíduo de uma típica função Cobb-Douglas, tal que o produto seja função do estoque de capital e do número total de trabalhadores.<sup>1</sup> Vale destacar que a participação das exporta-

<sup>1</sup> Como propõe Hall & Jones (1999), assumiu-se que a elasticidade do capital é igual a 1/3, e



Fonte: Elaboração dos autores.

**Figura 1:** Evolução trimestral do log da  $PTF_t$  (linha cheia) e do índice  $EX_t$  (linha tracejada)

ções de bens do país no total do valor exportado pelo mundo (EX), além de ser uma boa medida para caracterizar a trajetória de expansão das exportações, pode também ser considerada proxy para a produtividade do setor exportador.<sup>2</sup> Isso vem a significar, por exemplo, que aumentos da participação das exportações de um país no total exportado pelo mundo indicam a existência de alguma contrapartida favorável na produtividade do setor exportador, independentemente do sentido de orientação de causalidade em que isso venha a acontecer.<sup>3</sup>

No contexto aqui proposto, se houver indício de cointegração e de que a dinâmica de ajustamento se processa pela PTF, conclui-se que o movimento de causalidade de longo prazo se processa das exportações para a produtividade total dos fatores. Contudo, o vetor de correção de erros estimado irá impli-

---

com base no trabalho de Issler et al. (2004), foi considerado que o progresso tecnológico do país evolui exogenamente, de acordo com o modelo de Solow padrão. O estoque de capital inicial foi calculado tomando como base o procedimento de Hall & Jones (1999), considerando que o consumo de capital fixo foi utilizado como o montante depreciado a cada trimestre. Os dados foram obtidos do Fundo Monetário Internacional. Mais detalhes sobre o cálculo da PTF serão destacados adiante na seção 5 deste artigo.

<sup>2</sup>Essa suposição torna-se factível, uma vez que esse indicador, de modo bem abrangente, acaba por refletir os mecanismos econômicos convencionalmente propostos pela literatura que associa o papel das exportações das nações à sua produtividade. Pode-se citar a ocorrência de exploração de economias de escala, a melhora relativa da eficiência técnica, a melhoria dos termos de troca, dentre outros.

<sup>3</sup>A ideia principal para utilização desse índice é de justamente captar os movimentos antagônicos das exportações do país em relação ao total exportado pelo mundo. Por exemplo, quando há uma contrapartida favorável no índice, mesmo que seja via denominador, vem a significar que as exportações do país ainda se mantêm competitiva no comércio internacional. O país pode não ter sido influenciado pela tendência mundial de queda das exportações.

car, nesse caso, que a produtividade total dos fatores, já descontada dos efeitos do indicador de competitividade das exportações (PTFD), é determinada como resíduo da relação de longo prazo. Portanto, sua evolução torna-se meramente reativa à dinâmica de ajustamento ao equilíbrio deste sistema. Dada a conjunção de vários fatores que atuam sobre a PTF, essa restrição torna-se pouco factível, principalmente para países com setor não exportador amplo.

Via de regra, a literatura que faz o uso dessa abordagem empírica não leva em conta esse aspecto, pois grande parte das análises que testam a hipótese de orientação de crescimento via exportações se concentra em países em desenvolvimento que apresentam, em geral, economias com mercado doméstico pequeno e matriz produtiva preponderantemente voltada para fora.<sup>4</sup> Entretanto, quando se analisa países como os Estados Unidos e Japão, como propõe o artigo de Marin (1992), em que há mercado interno relativamente amplo e forte presença de setores com bens não comercializáveis, o modelo linear torna-se pouco factível.

Portanto, o que se faz necessário é relaxar a suposição imposta pelo vetor de correção de erros convencional de que, necessariamente, as exportações sempre restringem potenciais choques positivos de produtividade, não relacionados com o padrão de competitividade da atividade exportadora, fazendo com que a PTFD seja puramente reativa. Sendo assim, o que se sugere neste artigo é a hipótese de que essa dinâmica de ajustamento é não linear, tal que haja alternância entre regimes definidos tanto pelos desvios ao equilíbrio de longo prazo, como pelos regimes de retorno, que pode se processar de dois modos: via indicador de competitividade das exportações ou pela própria PTF.

Quais os motivos de ordem teórica para a existência desse padrão não linear? Para responder a essa questão, assume-se que a economia é dividida em dois principais setores. O primeiro agrega as firmas ligadas ao setor exportador, todavia, não exige a possibilidade de atuação no mercado doméstico. Já o segundo setor agrega as firmas não exportadoras com produção orientada exclusivamente ao mercado doméstico. Além disso, nessa economia, há possibilidade de *spillovers* tecnológicos entre os setores, possibilitando a existência de uma relação de longo prazo entre as produtividades setoriais e, portanto, a relação de longo prazo entre a PTF da economia e a produtividade do setor exportador. Como a análise aqui é dinâmica, assume-se adicionalmente que haveria processo de ajustamento no caso de eventuais desequilíbrios nessa relação. A questão, portanto, é identificar qual produtividade será responsável pelo ajuste, e isso vai estar diretamente relacionado às características de cada economia.

Nesse contexto, suponha uma economia com razão produto por trabalhador elevada, em que há mercado doméstico bastante amplo, com predominância da produção no segundo setor e baixa participação do primeiro setor no PIB. A ocorrência de choques positivos de produtividade, oriundos do setor não exportador, por exemplo, promoverá desvio em relação ao equilíbrio inicial. Como consequência, espera-se que o setor exportador seja responsável pelo restabelecimento do equilíbrio, uma vez que o aumento de produtividade do setor voltado ao mercado doméstico promove aumento da PTF que, por conseguinte, gera externalidades positivas à produtividade do setor ex-

---

<sup>4</sup>Em economias menores, torna-se mais provável que a escala de eficiência mínima seja maior do que o tamanho do mercado doméstico, indicando a necessidade de voltar parte de sua produção ao comércio internacional.

portador, elevando sua competitividade no comércio internacional. Ou seja, nesse regime, a produtividade total dos fatores orienta a competitividade das exportações.

Por outro lado, a ocorrência de choques negativos no setor não exportador também promoverá desvio em relação ao equilíbrio inicial, fazendo a PTF decrescer. Como a produtividade do primeiro setor responde à dinâmica competitiva no mercado internacional, há possibilidade desse choque não propiciar declínio em sua produtividade e, conseqüentemente, no nível de competitividade no comércio internacional. Como tendência de curto prazo, espera-se que o ajustamento se processe pelo próprio setor orientado para o mercado doméstico. O setor exportador, portanto, agiria como sustentador da PTF, revertendo a trajetória de queda. Isso poderia estar relacionado, por exemplo, ao aprofundamento da geração de externalidades positivas do primeiro para o segundo setor, tal como os termos de troca poderiam se tornar mais favoráveis às exportações, ou até mesmo, um deslocamento de firmas que passariam a comercializar seus produtos internacionalmente. Vale destacar que a velocidade de ajustamento aqui evidenciada dependerá de quão promissor é o mecanismo de catch-up das externalidades entre os setores.

Suponha agora uma economia que também possui relação produto por trabalhador elevada, mas com mercado doméstico menor e com boa parte de sua produção orientada para fora, conduzindo-o então à participação elevada das exportações no PIB. O mecanismo de reversão ao choque negativo de produtividade, orientada pelo setor exportador, não seria muito diferente da economia descrita anteriormente. Entretanto, quando há choques positivos de produtividade advindos do setor não exportador, não necessariamente implicará que a produtividade do setor exportador reaja endogenamente como se espera em países com mercado doméstico amplo, uma vez que em economias menores é possível que a produtividade do setor exportador possua menor sensibilidade de resposta à evolução da produtividade do setor orientado domesticamente. Poderia existir, por exemplo, no setor exportador mecanismo de desincentivo de captura de externalidades oriundas do setor não exportador. A sobrevivência de uma inovação, advinda do setor não exportador, é posta em xeque se o setor exportador não apropriá-la, tal como o custo de viabilizá-la pode ser muito superior aos benefícios a serem auferidos no comércio internacional. Portanto, haveria uma dissipação desse choque e o restabelecimento do equilíbrio ocorreria por meio do segundo setor, e, por conseguinte, da própria PTF. Conseqüentemente, nessa economia, o setor exportador agiria de modo a restringir a elevação da produtividade do país.

Toda a dinâmica de ajustamento, portanto, vai estar ligada ao regime de crescimento em que a PTF e a PTFD se encontram, seja em uma trajetória de crescimento ou de queda. Destacando que quanto maior o mecanismo de catch-up de externalidades entre os setores, maior a velocidade de ajuste ao equilíbrio. A Tabela 1, a seguir, apresenta a dinâmica não linear de desvio e de retorno ao equilíbrio de longo prazo entre a PTF e o indicador de competitividade das exportações, considerada aqui proxy para a produtividade do setor exportador.

Existem quatro diferentes situações em que o ajustamento está ou não habilitado, a depender das características de cada país. Considerando  $d > e$ , os regimes 2 e 3 refletem, respectivamente, os mecanismos de restrição e sustentação da PTF, caso seja esta a variável responsável pelo ajuste ao equilíbrio. Por sua vez, o regime 1 indica o desvio negativo à trajetória de longo prazo,

**Tabela 1:** Mecanismo não linear de desvio e de retorno ao equilíbrio de longo prazo em caso de choques sobre a PTF

Regime	$\sum \Delta PTF_{t-d}$ (choque)	$\sum \Delta PTF_{t-e}$	Movimento	Ajuste via $\Delta PTF_{t-1}$	Ajuste via $\Delta EX_{t-1}$	Mecanismos de Ajuste
1	< 0	< 0	Desvio ao Equilíbrio de Longo Prazo (< 0)	Não	Não	-
2	> 0	< 0	Retorno ao Equilíbrio de Longo Prazo (< 0)	Não Sim	Sim Não	Aumento de EX em resposta à elevação da PTF EX restringe aumento da PTF
3	< 0	> 0	Retorno ao Equilíbrio de Longo Prazo (> 0)	Sim Não	Não Sim	EX reverte queda da PTF Diminuição de EX em resposta à queda da PTF
4	> 0	> 0	Desvio ao Equilíbrio de Longo Prazo (> 0)	Não	Não	-

em que há um processo de decrescimento da PTF acumulada sucedido do declínio acumulado da PTFD, enquanto que o regime 4 indica o movimento inverso de perturbação do sistema dinâmico, com ascensão do acumulado da PTF sucedido de trajetória positiva do acumulado da PTFD.

Nessas circunstâncias, a família de modelos econométricos autorregressivos com transição suave é considerada aqui a mais apropriada para caracterizar essa dinâmica, uma vez que possibilita incorporar os quatro tipos regimes de ajustamento ao equilíbrio de longo prazo de acordo com a evolução de  $\sum \Delta PTF_{t-d}$  e  $\sum \Delta PTFD_{t-e}$ , descritos na Tabela 1.

### 3 O modelo MR-STVEC

Seguindo a proposta de Dijik et al. (2002) para modelos multivariados da família STAR, o modelo MR-STVEC com  $p$  defasagens e quatro regimes é especificado como:

$$\begin{aligned} \Delta x_t = & \left( \mu_1 + \alpha_1 z_{t-1} + \sum_{j=1}^p \Phi_{1,j} \Delta x_{t-j} \right) \left( 1 - F(s_{F_{t-d}}) \right) \left[ 1 - G(s_{G_{t-e}}) \right] + \\ & \left( \mu_2 + \alpha_2 z_{t-1} + \sum_{j=1}^p \Phi_{2,j} \Delta x_{t-j} \right) F(s_{F_{t-d}}) \left[ 1 - G(s_{G_{t-e}}) \right] + \\ & \left( \mu_3 + \alpha_3 z_{t-1} + \sum_{j=1}^p \Phi_{3,j} \Delta x_{t-j} \right) \left( 1 - F(s_{F_{t-d}}) \right) G(s_{G_{t-e}}) + \\ & \left( \mu_4 + \alpha_4 z_{t-1} + \sum_{j=1}^p \Phi_{4,j} \Delta x_{t-j} \right) F(s_{F_{t-d}}) G(s_{G_{t-e}}) + \varepsilon_t \end{aligned} \quad (1)$$

tal que,  $x_t = (x_{1t}, x_{2t}, \dots, x_{kt})'$  é um vetor ( $k \times 1$ ) séries de tempo;  $\Delta$  é o operador da primeira diferença;  $\mu_i$  para  $i = 1, 2, 3, 4$  são vetores ( $k \times 1$ ) de constantes;  $\alpha_i$  são matrizes ( $k \times r$ ) de coeficientes de ajustamento, tal que  $r$  é o número de vetores de cointegração;  $z_t = \beta' x_t$ , tal que,  $\beta$  é matriz ( $k \times r$ ) das relações lineares de longo prazo;  $\Phi_{i,j}$  para  $i = 1, 2, 3, 4$  e  $j = 1, \dots, p$  são matrizes dos coeficientes das defasagens; e  $\varepsilon_t$  é o vetor de ruídos brancos com média 0 e variância  $\Sigma$ . As funções  $F(s_{F_{t-d}})$  e  $G(s_{G_{t-e}})$  são logísticas<sup>5</sup> e limitadas entre 0 e 1, de acordo com Anderson & Teräsvirta (1992), Teräsvirta (1994) e Dijik et al. (2002), portanto,

$$F(s_{F_{t-d}}) = \left( 1 + \exp \left( \frac{-\gamma_F (s_{F_{t-d}} - c_F)}{\hat{\sigma}_{s_F}} \right) \right)^{-1} \quad (2)$$

$$G(s_{G_{t-e}}) = \left( 1 + \exp \left( \frac{-\gamma_G (s_{G_{t-e}} - c_G)}{\hat{\sigma}_{s_G}} \right) \right)^{-1} \quad (3)$$

tal que,  $s_{F_t}$  e  $s_{G_t}$  são as variáveis de transição,  $\gamma > 0$  é o parâmetro de suavidade (*smooth parameter*);  $d$  e  $e$  são os parâmetros de atraso (*delay parameters*) das funções de transição,  $\hat{\sigma}_{s_F}$  e  $\hat{\sigma}_{s_G}$  são os desvios padrões de  $s_{F_t}$  e  $s_{G_t}$ , respectivamente;  $c_F$  e  $c_G$  são os respectivos limiares (*thresholds*) de cada função

<sup>5</sup>Optou-se por utilizar a função logística uma vez que possibilita alternância entre dois tipos de regimes com características distintas, diferentemente da função exponencial que considera os valores extremos da série como único regime. Portanto, os desvios ao regime normal de crescimento representam apenas um regime.

de transição e indicam o momento em que ocorre a mudança de regime.<sup>6</sup> Portanto, o modelo possui relação de longo prazo linear, mas dinâmica de retorno ao equilíbrio não linear.<sup>7</sup> Assume-se, adicionalmente, que o modelo é globalmente estacionário, mas que pode apresentar regimes com trajetória de forte persistência. Logo, para avaliar se de fato as séries podem ser descritas de acordo com o modelo de múltiplos regimes, é necessário realizar inicialmente testes de especificação para a estrutura multivariada com duas funções logísticas.

#### 4 Especificação do Modelo MR-STVEC

A necessidade de uma estratégia coerente de especificação é requerida para avaliar se as séries podem ser caracterizadas adequadamente pelo modelo (1). Considera-se nesse processo, a escolha das variáveis de transição e a identificação dos parâmetros de atraso  $d$  ou  $e$ . Caso não haja indícios de que o melhor modelo é aquele descrito por (1), avalia-se se as séries podem ser caracterizadas por vetor de correção de erros com apenas uma função de transição (STVEC), como aplicado em Rothman et al. (2001), ou mesmo um VEC convencional. Nesse sentido, dois procedimentos serão executados para a decisão do modelo mais apropriado para sistema com duas equações.

O primeiro procedimento recomendado por Granger & Teräsvirta (1993), recorrentemente aplicado na literatura dos modelos da família STAR, é denominado específico para o geral. Consiste em realizar testes de especificação, a começar pela estimação do modelo de vetor de correção de erros mais simples, no caso o VEC, passando pelo STVEC, até o mais complexo, o MR-STVEC. O segundo procedimento, denominado específico para o geral para o específico, foi proposto em Lundbergh et al. (2003) e consiste em avaliar diretamente se o modelo MR-STVEC é mais apropriado do que o VEC e a partir daí é avaliado se a especificação STVEC com diferentes variáveis de transição é isoladamente apropriada para caracterizar as séries em questão.

Para a análise e melhor entendimento dos procedimentos aqui adotados, convém reparametrizar o modelo descrito em (1), da seguinte maneira:

$$\begin{aligned} \Delta x_t = & \tilde{\mu}_1 + \tilde{\alpha}_1 z_{t-1} + \sum_{j=1}^p \tilde{\Phi}_{1,j} \Delta x_{t-j} + \left( \tilde{\mu}_2 + \tilde{\alpha}_2 z_{t-1} + \sum_{j=1}^p \tilde{\Phi}_{2,j} \Delta x_{t-j} \right) F^*(s_{F_{t-d}}) + \\ & \left( \tilde{\mu}_3 + \tilde{\alpha}_3 z_{t-1} + \sum_{j=1}^p \tilde{\Phi}_{3,j} \Delta x_{t-j} \right) G^*(s_{G_{t-e}}) + \\ & \left( \tilde{\mu}_4 + \tilde{\alpha}_4 z_{t-1} + \sum_{j=1}^p \tilde{\Phi}_{4,j} \Delta x_{t-j} \right) F^*(s_{F_{t-d}}) G^*(s_{G_{t-e}}) + \varepsilon_t \end{aligned} \quad (4)$$

tal que,  $\tilde{\mu}_1 = \mu_1, \tilde{\mu}_2 = \mu_2 - \mu_1, \tilde{\mu}_3 = \mu_3 - \mu_1, \tilde{\mu}_4 = \mu_1 + \mu_4 - \mu_3 - \mu_2, \tilde{\alpha}_1 = \alpha_1, \tilde{\alpha}_2 = \alpha_2 - \alpha_1, \tilde{\alpha}_3 = \alpha_3 - \alpha_1, \tilde{\alpha}_4 = \alpha_1 + \alpha_4 - \alpha_3 - \alpha_2, \tilde{\Phi}_{1,j} = \Phi_{1,j}, \tilde{\Phi}_{2,j} = \Phi_{2,j} - \Phi_{1,j}, \tilde{\Phi}_{3,j} = \Phi_{3,j} - \Phi_{1,j}$  e  $\tilde{\Phi}_{4,j} = \Phi_{1,j} + \Phi_{4,j} - \Phi_{3,j} - \Phi_{2,j}$ . Além do mais,  $F^*(s_{F_{t-d}}) = F(s_{F_{t-d}}) - \frac{1}{2}$  e  $G^*(s_{G_{t-e}}) = G(s_{G_{t-e}}) - \frac{1}{2}$ . A reparametrização em (4) facilita enunciar as

<sup>6</sup>Se  $\gamma \rightarrow \infty$ , a função de transição possui transição discreta entre os regimes tal como a família dos modelos TAR (*Threshold Autorregressive*).

<sup>7</sup>O MR-STVEC pode ser considerado uma extensão do modelo de Rothman et al. (2001). Com essa estrutura, nos extremos 0 e 1 das funções de transição, o modelo assume 4 diferentes tipos de regimes.

hipóteses nulas de linearidade  $H_0 : \gamma_F = \gamma_G = 0$ ,  $H_0 : \gamma_F = 0$  e  $H_0 : \gamma_G = 0$  a serem testadas em cada procedimento. Entretanto, sob as hipóteses nulas enunciadas o MR-STVEC possui parâmetros não identificáveis. A literatura, neste caso, recomenda a realização da expansão de Taylor de primeira ordem ao redor de  $\gamma_F = \gamma_G = 0$ . Sendo assim, (4) pode ainda ser descrita como:

$$\begin{aligned} \Delta x_t = & \mu_1^* + \alpha_1^* z_{t-1} + \sum_{j=1}^p \Phi_{1,j}^* \Delta x_{t-j} + \left( \mu_2^* + \alpha_2^* z_{t-1} + \sum_{j=1}^p \Phi_{2,j}^* \Delta x_{t-j} \right) s_{F_{t-d}} + \\ & \left( \mu_3^* + \alpha_3^* z_{t-1} + \sum_{j=1}^p \Phi_{3,j}^* \Delta x_{t-j} \right) s_{G_{t-e}} + \\ & \left( \mu_4^* + \alpha_4^* z_{t-1} + \sum_{j=1}^p \Phi_{4,j}^* \Delta x_{t-j} \right) s_{F_{t-d}} s_{G_{t-e}} + R(\gamma_F, \gamma_G) + \varepsilon_t \end{aligned} \quad (5)$$

Os coeficientes a serem estimados  $\mu_i^*$ ,  $\alpha_i^*$  e  $\Phi_{i,j}^*$  para  $i = 1, 2, 3, 4$  e  $j = 1, \dots, p$  são combinações lineares dos coeficientes  $\tilde{\mu}_i$ ,  $\tilde{\alpha}_i$  e  $\tilde{\Phi}_{i,j}$  em (4). O termo  $R(\gamma_F, \gamma_G)$  representa o resto da expansão de Taylor de primeira ordem. Sob a hipótese nula esse termo é igual a zero, não trazendo prejuízo à distribuição do teste. A partir de (5) é possível executar os dois procedimentos aqui indicados para avaliar se a especificação em (1) é adequada.

#### 4.1 Procedimento “Específico para o Geral”

Inicia-se esse procedimento com o teste de linearidade proposto em Teräsvirta (1994), adaptado para sistema de duas equações, onde se avalia a hipótese nula de linearidade  $H_0 : \gamma_F = 0$  contra a hipótese alternativa  $H_0 : \gamma_F > 0$ . Tomando como base a especificação (4), assumindo-se a princípio que  $\tilde{\mu}_3 = \tilde{\mu}_4 = 0$ ,  $\tilde{\alpha}_3 = \tilde{\alpha}_4 = 0$  e  $\tilde{\Phi}_{3,j} = \tilde{\Phi}_{4,j} = 0$ , é avaliado se o modelo com apenas uma função de transição  $F(s_{F_{t-d}}, \gamma_F, c_F)$  é o mais apropriado do que o modelo linear. Como este modelo possui parâmetros não identificáveis sob a hipótese nula  $H_0 : \gamma_F = 0$ , realiza-se expansão de Taylor de primeira ordem ao redor de  $\gamma_F = 0$ . O resultado dessa expansão pode ser descrita por (5), assumindo-se que  $\mu_i^* = \alpha_i^* = \Phi_{i,j}^* = 0$ , para  $i = 3, 4$ . O teste de razão de verossimilhança (LR) é apropriado uma vez que se trata de sistema com mais de uma equação.

Refutada a hipótese nula de linearidade, estima-se modelo com uma função de transição ( $STVEC_F$ ) por mínimos quadrados não lineares. Em seguida, executa-se novo teste LR tomando como base o “*remaining nonlinearity test*” proposto por Eitrheim & Teräsvirta (1996) e discutido em Dijik et al. (2002), modificado para a estrutura de sistema de mais de uma equação. Refutada a hipótese de não linearidade adicional, há evidências de que as séries podem ser caracterizadas por (1).

As etapas detalhadas para a implementação dos testes são apresentadas a seguir.

- A. Estimar o modelo VEC linear e reportar os resíduos de cada uma das equações.
- B. Regredir os resíduos do passo anterior sobre a constante,  $z_{t-1}$  e  $\Delta x_{t-j}$  para  $j = 1, \dots, p$  e computar a matriz de variância-covariância  $\Sigma_{11}$ .

- C. Regredir os resíduos do VEC linear sobre a constante,  $z_{t-1}, z_{t-1}S_{F_{t-d}}$  e  $\Delta x_{t-j}S_{F_{t-d}}$  para  $j = 1, \dots, p$  e computar a matriz de variância-covariância  $\Sigma_{12}$ .
- D. Calcular a estatística  $LR_1 = (T - c) \{\log(\Sigma_{11}) - \log(\Sigma_{12})\}$  distribuída de acordo com uma  $\chi^2$  com grau de liberdade igual ao número de restrições do sistema, em que  $T$  é o número de observações usadas e  $c$  (correção de Sims para pequenas amostras) é igual ao número de parâmetros estimados em cada equação do sistema não restrito.
- E. Caso se refute a hipótese nula de linearidade, a pelo menos 10% de significância, estima-se o sistema de equações por mínimos quadrados não lineares multivariado. Como em Teräsvirta (1994) e Lundbergh et al. (2003), o valor do parâmetro de atraso  $d$  e da variável de transição  $S_{F_{t-d}}$  são escolhidos com base no menor p-valor do teste de linearidade.<sup>8</sup> Define-se o modelo  $STVEC_F$  estimado através da seguinte especificação:

$$\begin{aligned} \Delta x_t &= H(z_{t-1}, \Delta x_{t-j}, S_{F_{t-d}}; \hat{\Psi}) + \hat{\varepsilon}_t \\ &= H(z_{t-1}, \Delta x_{t-j}, S_{F_{t-d}}; \hat{\mu}_1, \hat{\mu}_2, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\Phi}_{1,j}, \hat{\Phi}_{2,j}, \hat{\gamma}_F, \hat{c}_F) + \hat{\varepsilon}_t \end{aligned} \quad (6)$$

tal que,

$$\begin{aligned} H(z_{t-1}, \Delta x_{t-j}, S_{F_{t-d}}; \hat{\Psi}) &= \hat{\mu}_1 + \hat{\alpha}_1 z_{t-1} + \sum_{j=1}^p \hat{\Phi}_{1,j} \Delta x_{t-j} + \hat{\varepsilon}_t + \\ &\left( \hat{\mu}_2 + \hat{\alpha}_2 z_{t-1} + \sum_{j=1}^p \hat{\Phi}_{2,j} \Delta x_{t-j} \right) F(S_{F_{t-d}}, \hat{\gamma}_F, \hat{c}_F) \end{aligned} \quad (7)$$

- F. Regredir o resíduo do modelo  $STVEC_F$  sobre  $\nabla H(z_{t-1}, \Delta x_{t-j}, S_{F_{t-d}}; \hat{\Psi}) = \frac{\partial H(z_{t-1}, \Delta x_{t-j}, S_{F_{t-d}}; \hat{\Psi})}{\partial \hat{\Psi}}$ ,<sup>9</sup> para assim obter a matriz de variância-covariância do modelo restrito  $\Sigma_{21}$ .
- G. Regredir cada resíduo do modelo  $STVEC_F$  sobre  $\nabla H(z_{t-1}, \Delta x_{t-j}, S_{F_{t-d}}; \hat{\Psi}) = \frac{\partial H(z_{t-1}, \Delta x_{t-j}, S_{F_{t-d}}; \hat{\Psi})}{\partial \hat{\Psi}}$  e sobre a expansão de Taylor de primeira ordem da função  $G(s_{G_{t-e}}, \gamma_G, c_G)$ , em torno de  $\gamma_G = 0$ , ou seja,  $z_{t-1}S_{G_{t-e}}, \Delta x_{t-j}S_{G_{t-e}}, z_{t-1}F(S_{F_{t-d}}, \hat{\gamma}_F, \hat{c}_F)S_{G_{t-e}}$  e  $\Delta x_{t-j}F(S_{F_{t-d}}, \hat{\gamma}_F, \hat{c}_F)S_{G_{t-e}}$  para  $j = 1, \dots, p$  e obter a matriz de variância-covariância do modelo não restrito  $\Sigma_{22}$ .<sup>10</sup>
- H. Calcular a estatística  $LR_2 = (T - c) \{\log(\Sigma_{21}) - \log(\Sigma_{22})\}$  distribuída de acordo com uma  $\chi^2$  com grau de liberdade igual ao número de restrições do sistema, em que  $T$  é o número de observações usadas e  $c$  (correção de Sims para pequenas amostras) é igual ao número de parâmetros estimados em cada equação do sistema não restrito. Utiliza-se novamente

<sup>8</sup> Ver detalhes em Teräsvirta (1994, p. 211) e Lundbergh et al. (2003, p. 107 e 113).

<sup>9</sup> Para mais detalhes sobre a derivada parcial com relação aos coeficientes estimados do STVEC, consultar Apêndice A.

<sup>10</sup> O artigo de Dijk & Franses (1999) apresenta procedimento semelhante para o caso univariado.

como critério o menor p-valor para a escolha da variável de transição  $s_{G_{t-e}}$  e do parâmetro de atraso  $e$ .

Se a hipótese nula é refutada, então o MR-STVEC pode ser estimado por mínimos quadrados não lineares. Encerra-se, portanto, o procedimento “específico para o geral”.

#### 4.2 Procedimento “Específico para o Geral para o Específico”

Neste procedimento, Lundbergh et al. (2003) propõem testar diretamente a especificação linear contra a especificação do MR-STVEC. Se a hipótese nula de linearidade for rejeitada, testam-se duas sub-hipóteses em que se avalia se os modelos  $STVEC_F$  e  $STVEC_G$  são individualmente necessários para a caracterização das séries. Se essas duas hipóteses forem refutadas, temos indícios de que as séries podem ser descritas como um MR-STVEC. Os testes são realizados por meio da estimação do modelo (5). Os passos necessários para aplicar o procedimento específico para o geral para o específico são apresentados a seguir:

- A. Estimar o modelo VEC linear e reportar os resíduos de cada uma das equações.
- B. Regredir o vetor de resíduos do passo anterior sobre  $z_{t-1}$  e  $\Delta x_{t-j}$  para  $j = 1, \dots, p$  e computar a matriz de variância-covariância  $\tilde{\Sigma}_{11}$ .
- C. Regredir o vetor de resíduos do VEC linear sobre  $z_{t-1}$ ,  $\Delta x_{t-j}$ ,  $z_{t-1}s_{F_{t-d}}$ ,  $\Delta x_{t-j}s_{F_{t-d}}$ ,  $z_{t-1}s_{G_{t-e}}$ ,  $\Delta x_{t-j}s_{G_{t-e}}$ ,  $z_{t-1}s_{F_{t-d}}s_{G_{t-e}}$  e  $\Delta x_{t-j}s_{F_{t-d}}s_{G_{t-e}}$  para  $j = 1, \dots, p$  e computar  $\tilde{\Sigma}_{12}$ .
- D. Calcular a estatística  $LR'_1 = (T - c) \{ \log(\tilde{\Sigma}_{11}) - \log(\tilde{\Sigma}_{12}) \}$  que é distribuída de acordo com uma  $\chi^2$  com grau de liberdade igual ao número de restrições do sistema, em que  $T$  é o número de observações usadas e  $c$  (correção de Sims para pequenas amostras) é igual ao número de parâmetros estimados em cada equação do sistema não restrito.
- E. Se rejeitarmos a hipótese nula  $H_0^{MR-STVEC}$  de que o modelo é um VEC, a pelo menos 10% de significância, realiza-se dois testes de hipótese auxiliares adicionais  $H_0^{STVEC_F}$  e  $H_0^{STVEC_G}$  que avaliam se os modelos  $STVEC_F$  e  $STVEC_G$  de modo isolado são bons para caracterizar as séries de tempo. Ou seja, regredir, respectivamente, o vetor de resíduos estimados do VEC linear sobre  $z_{t-1}$ ,  $\Delta x_{t-j}$ ,  $z_{t-1}s_{G_{t-e}}$  e  $\Delta x_{t-j}s_{G_{t-e}}$  para  $j = 1, \dots, p$  e computar a matriz de variância e covariância  $\tilde{\Sigma}_{13}$ , assim como, regredir o resíduos estimados obtidos do modelo VEC linear sobre  $z_{t-1}$ ,  $\Delta x_{t-j}$ ,  $z_{t-1}s_{F_{t-d}}$  e  $\Delta x_{t-j}s_{F_{t-d}}$  para  $j = 1, \dots, p$  e computar a matriz de variância e covariância  $\tilde{\Sigma}_{14}$ .
- F. Calcular as estatísticas  $LR'_3 = (T - c) \{ \log(\tilde{\Sigma}_{13}) - \log(\tilde{\Sigma}_{12}) \}$  e  $LR'_4 = (T - c) \{ \log(\tilde{\Sigma}_{14}) - \log(\tilde{\Sigma}_{12}) \}$  distribuídas de acordo com  $\chi^2$  uma com grau de liberdade igual ao número de restrições do sistema, em que  $T$  é o número de observações usadas e  $c$  (correção de Sims para pequenas amostras) é igual ao número de parâmetros estimados em cada equação do sistema não restrito.

Após a execução dos passos apresentados, se a hipótese nula  $H_0^{MR-STVEC}$  é rejeitada, assim como  $H_0^{STVEC_F}$  e  $H_0^{STVEC_G}$  são rejeitadas, há indícios de que individualmente os modelos  $STVEC_F$  e  $STVEC_G$  são necessários para caracterizar as séries de tempo, portanto, estima-se o  $MR-STVEC$ . No caso de a hipótese nula  $H_0^{STVEC_F}$  ser rejeitada, mas  $H_0^{STVEC_G}$  não ser rejeitada, significa que individualmente o  $STVEC_G$  não é apto para caracterização, portanto, somente o modelo  $STVEC_F$  deve ser estimado. O mesmo raciocínio é válido quando  $H_0^{STVEC_G}$  for rejeitada e  $H_0^{STVEC_F}$  não for rejeitada.<sup>11</sup>

## 5 Avaliação dos Resultados

### 5.1 Dados

Para este estudo foram obtidas as séries trimestrais e com ajuste sazonal do PIB, formação bruta de capital, total da força de trabalho e receita de exportações de bens em dólares, a preços correntes, para os Estados Unidos, Canadá, Japão e Alemanha, além da série do valor total exportado de bens no mundo em dólares, a preços correntes. As séries foram obtidas do banco de dados Internacional Financial Statistic do FMI e abrangem o período entre o primeiro trimestre de 1957 ao quarto trimestre de 2009, exceção feita ao caso alemão, em que as séries iniciam-se no primeiro trimestre de 1962. Somente para as séries de exportações de bens foi necessário realizar o ajuste sazonal.<sup>12</sup>

A série  $\{EX_t\}$  foi calculada por meio da razão entre as receitas de exportações nominais do país e o total de receitas do exportado pelo mundo. Para o computo da produtividade total dos fatores  $\{PTF_t\}$ , as séries do PIB e da formação bruta de capital foram deflacionadas para o ano de 2005, por meio do deflator do PIB de cada país, também obtido da base de dados do FMI. O estoque de capital foi estimado com base na proposta de Hall & Jones (1999). O montante inicial do estoque de capital foi obtido pelo cálculo da razão entre o valor inicial da série de formação bruta de capital ( $I_0$ ) e a soma entre a taxa geométrica de crescimento trimestral da formação bruta de capital nos 10 primeiros anos da série  $g$  e a taxa média de depreciação  $\delta$ , portanto,  $\frac{I_0}{(g+\delta)}$ . A base de dados do FMI também fornece o consumo de capital fixo, que após ser deflacionado, foi utilizada como montante depreciado em cada trimestre, tal como para a estimativa da taxa média de depreciação  $\delta$ . Por simplicidade, assumiu-se que o valor da elasticidade do capital é igual a 1/3 (Cooley & Prescott 1995, Hall & Jones 1999). Em vez de se assumir valor fixo para o coeficiente do progresso tecnológico exógeno  $\theta$ , optou-se por estimá-lo através de mínimos quadrados.

As variáveis de transição para o modelo MR-STVEC foram construídas através do cálculo do somatório da variação trimestral da produtividade total dos fatores ( $\Delta \log PTF_t$ ), variando de 1 a 8 trimestres ( $h_1 = 1, \dots, 8$ ), e da produtividade total dos fatores descontada ( $\Delta PTFD_t$ ), também variando de 1 a 8 trimestres ( $h_2 = 1, \dots, 8$ ), ou seja,

<sup>11</sup>O mecanismo não linear de ajustamento aqui proposto ainda não foi discutido pela literatura. Portanto, torna-se necessário avaliar o poder e tamanho do teste para ambos os procedimentos, levando-se em conta os possíveis mecanismos de correção de erros presente nesta análise. Tal exercício esta presente no Anexo C deste artigo disponível somente na versão online.

<sup>12</sup>O ajuste sazonal foi realizado com base no método Census X12.

$$s_{F_{t-d}} = \left( \sum_{h=1}^{h=h_1} \Delta \log PTF_{t-h+1} \right)_{t-d} \quad \text{ou} \quad \left( \sum_{h=1}^{h=h_2} \Delta \log PTF_{t-h+1} \right)_{t-d} \quad (8)$$

$$s_{G_{t-e}} = \left( \sum_{h=1}^{h=h_1} \Delta \log PTF_{t-h+1} \right)_{t-e} \quad \text{ou} \quad \left( \sum_{h=1}^{h=h_2} \Delta \log PTF_{t-h+1} \right)_{t-e} \quad (9)$$

Vale destacar que a  $PTFD_t$  é calculada como resíduo da relação de longo prazo entre o log da  $PTF_t$  e a razão das exportações  $EX_t$ .

## 5.2 Resultados dos Testes de Especificação

Os testes de especificação propostos foram realizados com base na estrutura do VEC linear estimado pelo método de Johansen (1988, 1991). O critério de informação de Akaike (AIC) foi utilizado para a escolha do número de defasagens do sistema, levando-se em conta também a ausência de autocorrelação nos resíduos de cada equação.<sup>13</sup> Para todos os casos, o teste traço de razão de verossimilhança indicou a existência de um vetor de cointegração, como aponta a Tabela B.1 no Apêndice B.

Os parâmetros de atraso (8) e (9) são dados por  $d = 1, \dots, d_{max}$  e  $e = 1, \dots, e_{max}$ , com  $d_{max} = e_{max} = 4$ . Uma vez que o procedimento “específico para o geral para o específico” envolve a realização do teste para o total de 1024 diferentes modelos estimados, a Tabela 2, apresenta somente os cinco menores p-valores ligados à hipótese  $H_0^{MR-STVEC}$  e os subtestes auxiliares para cada país aqui em análise.

Ainda em relação aos resultados da Tabela 2, considerou-se que o valor do parâmetro de atraso  $d$  é necessariamente superior ao parâmetro  $e$ . Lembrando que a escolha dos parâmetros de atraso e do valor acumulado das taxas de crescimento das variáveis de transição segue a proposta de Teräsvirta (1994), ou seja, selecionar os parâmetros que apresentarem o menor p-valor. Para todos os países, há indícios de que a hipótese nula  $H_0^{MR-STVEC}$  é rejeitada a pelo menos 1% de significância. Do mesmo modo, as duas sub-hipóteses  $H_0^{STVEC_F}$  e  $H_0^{STVEC_G}$  foram rejeitadas a níveis convencionais de significância. Logo, de acordo com o procedimento em questão, há fortes evidências de que as séries podem ser descritas de acordo com um MR-STVEC.

Resta verificar se o procedimento “específico para o geral” também conduz às mesmas conclusões. Repare que em grande parte dos casos,  $h_1$  e  $h_2$  apresentaram valores superiores a 1, conferindo a variável de transição valores acumulados da taxa de crescimento trimestral das séries que podem chegar próximos a 2 anos, como no caso alemão.

A coluna 4 da Tabela 3 apresenta os p-valores do teste do procedimento “específico para o geral”, cuja hipótese nula verifica se a especificação do modelo  $STVEC_F$  é mais adequada do que o modelo VEC linear. Concomitantemente, a coluna 8 apresenta os p-valores do teste em que se verifica se a especificação do modelo  $STVEC_G$  é mais adequada do que o modelo VEC linear.

O procedimento “específico para o geral” apresenta resultados compatíveis ao procedimento anterior. Observa-se que para todos os pares  $(h_1, d)$  destacados na Tabela 3, houve rejeição da hipótese nula quando associada a  $F(\cdot)$ , a pelo menos 5% de significância, com p-valores inferiores aos pares  $(h_2, e)$  correspondentes. Seguindo a estratégia de Lundbergh et al. (2003), faz-se

<sup>13</sup>Os resultados do teste Dickey-Fuller de raiz unitária, não reportados, indicaram que todos os processos  $\log PTF_t$  e de  $EX_t$  são integrados de ordem 1.

**Tabela 2:** p-valores do Teste L para o procedimento “específico para o geral para o específico”

Países	Variável de Transição F(.)		Variável de Transição G(.)		Hipóteses Nulas		
	$\left(\sum_{h=1}^{h=h_1} \Delta \log PTF_{t-h+1}\right)_{t-d}$		$\left(\sum_{h=1}^{h=h_2} \Delta \log PTFD_{t-h+1}\right)_{t-e}$		$H_0^{MR-SVTEC}$	$H_0^{STVECF}$	$H_0^{STVECG}$
	$h_1$	d	$h_2$	e			
Estados Unidos	2	4	2	1	0,000006	0,000002	0,000144
	2	4	6	3	0,000008	0,000003	0,000178
	2	4	7	2	0,000010	0,000001	0,000222
	3	4	2	1	0,000012	0,000003	0,000054
	5	4	5	1	0,000015	0,000044	0,000033
Canadá	3	3	1	2	0,000426	0,000172	0,002604
	2	4	1	2	0,000518	0,000216	0,006803
	2	4	2	1	0,000754	0,000208	0,009859
	3	3	6	2	0,000806	0,000425	0,002758
	3	3	7	1	0,001099	0,000139	0,003846
Japão	1	4	4	3	0,000001	0,000001	0,000121
	1	4	3	4	0,000003	0,000001	0,000299
	1	4	6	1	0,000033	0,000009	0,002669
	3	2	6	1	0,000040	0,000011	0,000394
	4	3	1	1	0,000041	0,000089	0,000006
Alemanha	2	2	7	1	0,000011	0,000000	0,002097
	2	2	1	1	0,000078	0,001003	0,035171
	2	2	2	1	0,000091	0,000018	0,039845
	2	2	8	1	0,000163	0,000005	0,015078
	2	2	5	1	0,000851	0,000046	0,185063

**Tabela 3:** p-valores do Teste LR para o procedimento “específico para o geral”

Países	Variável de Transição F(.)		Hipóteses Nulas associadas a F(.)		Variável de Transição G(.)		Hipóteses Nulas associadas a G(.)	
	$\left(\sum_{h=1}^{h=h_1} \log \Delta PTF_{t-h+1}\right)_{t-d}$		$H_0^{VEC}$	$H_0^{STVECF}$	$\left(\sum_{h=1}^{h=h_2} \log \Delta PTF_{t-h+1}\right)_{t-e}$		$H_0^{VEC}$	$H_0^{STVECG}$
	$h_1$	d			$h_2$	e		
Estados Unidos	2	4	0,00143	0,05830	2	1	0,10190	0,01771
	2	4	0,00143	0,05830	6	3	0,11249	0,02606
	2	4	0,00143	0,05830	7	2	0,27539	0,07192
	3	4	0,01188	0,01051	2	1	0,10190	0,05903
	5	4	0,02444	0,00751	5	1	0,01302	0,00685
Canadá	3	3	0,01057	0,20571	1	2	0,14581	0,04792
	2	4	0,00378	0,41775	1	2	0,14581	0,06495
	2	4	0,00378	0,41775	2	1	0,15705	0,01220
	3	3	0,01057	0,20571	6	2	0,15668	0,43239
	3	3	0,01057	0,20571	7	1	0,45580	0,43221
Japão	1	4	0,00023	0,00428	4	3	0,04406	0,00244
	1	4	0,00023	0,00428	3	4	0,14601	0,00519
	1	4	0,00023	0,00428	6	1	0,17794	0,00300
	3	2	0,00666	0,21324	6	1	0,17794	0,00775
	4	3	0,27109	0,01945	1	1	0,02423	0,03092
Alemanha	2	2	0,00001	0,15404	7	1	0,71918	0,01463
	2	2	0,00001	0,15404	1	1	0,00328	0,16036
	2	2	0,00001	0,15404	2	1	0,23594	0,10768
	2	2	0,00001	0,15404	8	1	0,67799	0,07020
	2	2	0,00001	0,15404	5	1	0,65335	0,51445

necessário estimar os modelos STVEC com a função de transição  $F(\cdot)$  e novamente realizar os testes de especificação, só que dessa vez, para avaliar se o modelo MR-STVEC é mais adequado do que o modelo STVEC. Os p-valores do segundo estágio dos testes para as funções de transição  $F(\cdot)$  e  $G(\cdot)$  já estimados são apresentados, respectivamente, nas colunas 5 e 9 da Tabela 3. Vale ainda destacar que, de maneira similar aos resultados das simulações por Monte Carlo, no Apêndice C, os testes  $LR$  sob as hipóteses nulas associadas a  $G(\cdot)$  apresentaram níveis de rejeição maiores quando comparado ao estágio anterior.

Por conseguinte, verifica-se que há evidências suficientes, com base na estratégia de especificação aqui delineada de que as séries do log da PTF e da razão das exportações podem ser descritas de acordo com um MR-STVEC. No caso alemão, não houve dificuldades para a escolha dos parâmetros. Os pares  $(h_1 = 2, d = 2)$  e  $(h_2 = 7, e = 1)$  propiciaram os menores p-valores no “procedimento específico para o geral para o específico” tal como nas duas hipóteses nulas dos estágios do procedimento “específico para o geral”. O mesmo se verifica para o Japão com os pares  $(h_1 = 1, d = 4)$  e  $(h_2 = 4, e = 3)$ . Já no caso canadense, devido ao melhor ajuste do modelo, observado pelo critério de informação  $AIC$ , selecionou-se as combinações  $(h_1 = 2, d = 4)$  e  $(h_2 = 2, e = 1)$ , que são compatíveis com os menores p-valores em cada estágio do procedimento “específico para o geral”. Somente para os Estados Unidos, optou-se por escolher os pares  $(h_1 = 3, d = 4)$  e  $(h_2 = 2, e = 1)$ , mesmo que não tenham sido as combinações com os menores p-valores para ambos os procedimentos, uma vez que os modelos MR-STVECs estimados para o par  $(h_1 = 2, d = 4)$  não apresentaram o melhor ajustamento aos dados, principalmente pela presença de heterocedasticidade condicional nos resíduos.

Portanto, a partir dos resultados da estratégia de especificação aqui proposta, recorre-se a estimação dos modelos VEC de múltiplos regimes com transição suave.

### 5.3 Análise dos Resultados

Além disso, o teste ARCH aplicado aos resíduos, ou seja, minimiza-se a expressão  $\sum_t u_t \Sigma u_t$ , tal que,  $u_t$  é o vetor de resíduos e  $\Sigma = E u_t' u_t$ . Um *grid* com os valores fixos dos coeficientes do parâmetro de suavidade é utilizado para escolha dos valores iniciais. Aqueles que engendrarem o menor log do determinante de  $\Sigma$  são utilizados para tanto. Como a proposta empírica discutida na seção 2 implica que a mudança de regimes ocorre de acordo com a alternância dos sinais da taxa de crescimento do acumulado do  $\log$  da  $PTF_t$  e da  $PTFD_t$ , assumiu-se que os valores iniciais dos limiares são iguais a zero.

Os resultados dos modelos estimados para os quatro países em análise são apresentados na Tabela 4.<sup>14</sup> Alguns pontos são importantes de serem ressaltados previamente à discussão dos resultados. Primeiramente, observa-se que os modelos não lineares se ajustam melhor às séries em destaque, como apontam os critérios de informação  $AIC$  e da razão entre os desvios padrões do modelo linear e não linear inferiores à unidade. Além disso, o teste  $ARCH$ , aplicado aos resíduos de cada equação, não traz indícios da existência de heterocedasticidade condicional, a pelo menos 5% de significância. Em segundo

<sup>14</sup>Para salvar espaço, os coeficientes das defasagens do MR-STVEC não foram reportados.

lugar, deve-se levar em conta que os quatro regimes, apresentados na Tabela 1, são caracterizados pelo MR-STVEC através da seguinte dinâmica:

- Regime 1: desvio negativo ao equilíbrio de longo prazo, ou seja,  $\left(\sum_{h=1}^{h=h_1} \Delta \log PTF_{t-h+1}\right)_{t-d} \ll 0$  e  $\left(\sum_{h=1}^{h=h_2} \Delta \log PTFD_{t-h+1}\right)_{t-e} \ll 0$ ;
- Regime 2: retorno ao equilíbrio de longo prazo depois de desvio positivo da PTF, ou seja, e  $\left(\sum_{h=1}^{h=h_1} \Delta \log PTF_{t-h+1}\right)_{t-d} \gg 0$  e  $\left(\sum_{h=1}^{h=h_2} \Delta \log PTFD_{t-h+1}\right)_{t-e} \ll 0$ ;
- Regime 3: retorno ao equilíbrio de longo prazo depois de desvio negativo da PTF, ou seja,  $\left(\sum_{h=1}^{h=h_1} \Delta \log PTF_{t-h+1}\right)_{t-d} \ll 0$  e  $\left(\sum_{h=1}^{h=h_2} \Delta \log PTFD_{t-h+1}\right)_{t-e} \gg 0$ ;
- Regime 4: desvio positivo ao equilíbrio de longo prazo, ou seja,  $\left(\sum_{h=1}^{h=h_1} \Delta \log PTF_{t-h+1}\right)_{t-d} \gg 0$  e  $\left(\sum_{h=1}^{h=h_2} \Delta \log PTFD_{t-h+1}\right)_{t-e} \gg 0$ .

Feitas estas considerações, a partir dos resultados da Tabela 4, nota-se que o modelo *VEC* linear estimado indica que, exceto no caso alemão, somente os coeficientes de ajustamento da equação da *PTF* são estatisticamente significantes a níveis convencionais. Como salientado previamente, para países como os Estados Unidos e Japão, tal resultado não parece ser plausível dado que, ao longo de quase toda a segunda metade do século XX, choques positivos de produtividade, oriundos de outros fatores não associados ao setor externo, são restringidos pelas exportações.

Por outro lado, os resultados do MR-STVEC apresentam outra perspectiva. Para todos os casos, esse mecanismo de ajustamento vem ocorrer no regime 3, quando a *PTF* dessas economias encontra-se em fase de baixo crescimento, seguido de movimento de recuperação da *PTFD*. Após o sistema sofrer choques negativos de produtividade, há uma dinâmica de reversão, significando a existência de mecanismo de sustentação da *PTF* orientado pelas exportações.

O MR-STVEC estimado para Japão e Alemanha evidenciam que, por meio das exportações, o ajustamento de curto prazo vem ocorrer no regime 2. Por conseguinte, quando o *log* da *PTF* encontra-se em movimento de ascensão, seguido de queda da *PTFD*, *EX* é a variável que responde positivamente. Logo, choques que promovam maior nível de produtividade, seja ele oriundo do setor não exportador, serão traduzidos em aumento do nível de participação das exportações. Interessante observar que para os Estados Unidos não há evidência de qualquer tipo de resposta do setor exportador ao aumento de produtividade da economia, uma vez que o coeficiente de ajustamento no regime 2 não é estatisticamente significativo a níveis convencionais. O que, portanto, leva ao questionamento sobre o quão sensível é o setor exportador aos choques positivos de produtividade.

O Canadá e a Alemanha são os países que destoam das outras duas economias pelo fato de possuírem coeficiente de ajustamento no regime 2 da equação do *log* da *PTF<sub>t</sub>* estatisticamente diferente de zero. Dessa forma, logo após a ocorrência de choques positivos de produtividade há trajetória de retorno ao patamar pré-choque. O que chama atenção, portanto, é que o nível de participação das exportações canadenses restringe possíveis choques positivos de produtividade. Esse aspecto pode ser explicado pelas características desse

**Tabela 4:** Resultados da Estimação do MR-STVEC

Modelo	Estados Unidos						Canadá						
	VEC			MR-STVEC			VEC			MR-STVEC			
Equação	n	$\Delta \log PTF_t$	$\Delta X_t$	n	$\Delta \log PTF_t$	$\Delta X_t$	n	$\Delta \log PTF_t$	$\Delta X_t$	n	$\Delta \log PTF_t$	$\Delta X_t$	
Regime 1	$\hat{\alpha}_1$	207	-0,0436***	0,0029	80	-0,0351	0,0029	207	-0,0449***	-0,0018	79	-0,0549	-0,0011
Regime 2	$\hat{\alpha}_2$				45	-0,0306	-0,0137				28	-0,0723**	-0,0067
Regime 3	$\hat{\alpha}_3$				43	-0,0933***	-0,0179				71	-0,0597**	-0,0053
Regime 4	$\hat{\alpha}_4$				37	-0,0121	0,0048				27	0,0271	0,0016
$\gamma_F /$	$\gamma_G$					9,0	17,0					2,0	20,0
$C_F /$	$C_G$					0,0057***	0,0056***					0,0078**	0,0030***
F(.)	$h_1$					3						2	
	$d$					4						4	
	$h_2$					2						2	
G(.)	$e$					1						1	
p-valor			0,01063	0,00002		0,21291	0,40092		0,00648	0,31177		0,70627	0,95386
ARCH(4)													
p-valor			0,89501	0,93154		0,72521	0,12068		0,99481	0,99584		0,72127	0,13927
A(4)													
$\hat{\sigma}_{MR-STVEC}$						0,92	0,91					0,83	0,85
$\hat{\sigma}_{VEC}$													
AIC			-4138,08			-4132,28			-4562,36			-4588,60	

Nota: ARCH(.) indica o teste de heterocedasticidade condicional sobre os resíduos de cada equação; A(.) indica o teste Q de autocorrelação serial de Ljung-Box para o modelo VEC e o teste proposto por Eitrheim & Teräsvirta (1996) para o modelo MR-STVEC. \* 10% de significância, \*\* 5% de significância, \*\*\* 1% de significância.

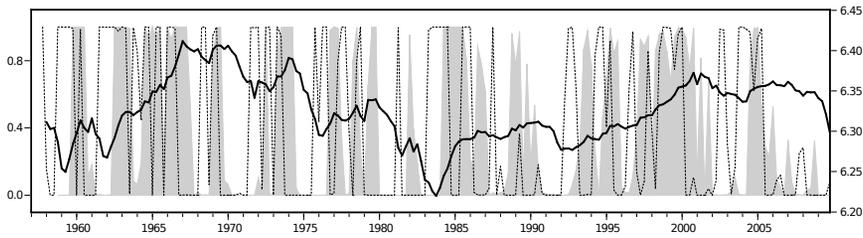
**Tabela 4:** Resultados da Estimação do MR-STVEC (continuação)

Modelo	Japão						Alemanha						
	VEC			MR-STVEC			VEC			MR-STVEC			
Equação	n	$\Delta \log PTF_t$	$\Delta X_t$	n	$\Delta \log PTF_t$	$\Delta X_t$	n	$\Delta \log PTF_t$	$\Delta X_t$	n	$\Delta \log PTF_t$	$\Delta X_t$	
Regime 1	$\hat{\alpha}_1$	207	-0,0405***	0,0005	66	-0,0121	-0,0043	187	-0,0615***	0,0122**	59	-0,0373	0,0197
Regime 2	$\hat{\alpha}_2$				42	-0,0473	0,0164**				21	-0,1233*	0,0456*
Regime 3	$\hat{\alpha}_3$				28	-0,0816**	0,0015				57	-0,0549*	0,0111
Regime 4	$\hat{\alpha}_4$				69	-0,0254	-0,0021				49	-0,0225	0,0182
$\gamma_F /$	$\gamma_G$				21,0	73,0					20,0	200,0	
$CF /$	$CG$				0,0000	0,0093***					0,0045***	0,0018	
F(.)	$h_1$				1						2		
	$d$				4						2		
G(.)	$h_2$				4						7		
	$e$				3						1		
p-valor		0,00758	0,53751		0,70422	0,36758		0,0982	0,1967		0,5988	0,4416	
ARCH(4)													
p-valor		0,98464	0,99891		0,12098	0,92408		0,32956	0,99083		0,9691	0,8435	
A(4)													
$\frac{\hat{\sigma}_{MR-STVEC}}{\hat{\sigma}_{VEC}}$					0,79	0,86					0,73	0,89	
AIC		-4186,14			-4216,59			-3704,40			-3740,90		

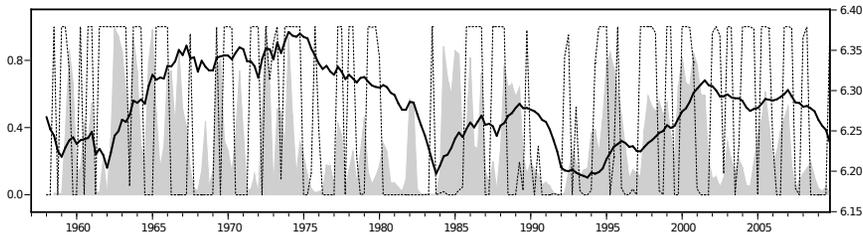
Nota: ARCH(.) indica o teste de heterocedasticidade condicional sobre os resíduos de cada equação; A(.) indica o teste Q de autocorrelação serial de Ljung-Box para o modelo VEC e o teste proposto por Eitrheim & Teräsvirta (1996) para o modelo MR-STVEC. \* 10% de significância, \*\* 5% de significância, \*\*\* 1% de significância.

país que possui mercado doméstico relativamente pequeno e alta participação das exportações no PIB. Já no caso alemão, para o mesmo regime 2, as exportações também são responsáveis pelo ajustamento de curto prazo com a seguinte característica: o coeficiente  $\hat{\alpha}_{21}$  (equação da  $PTF_t$ ) superior a  $\hat{\alpha}_{22}$  (equação de  $EX_t$ ). Portanto, choques positivos sobre a PTF fazem com que o mecanismo de restrição seja parcial, tal como, há contrapartida positiva na participação de suas exportações no total exportado mundialmente.

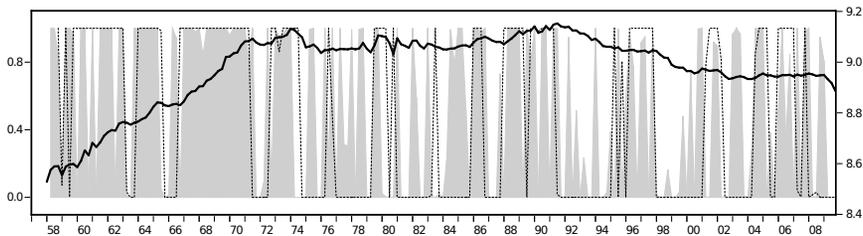
Para melhor compreensão a respeito do processo de ajustamento de curto prazo nos modelos MR-STVEC, as Figuras 2 a 5, a seguir, apresentam a evolução das funções de transição conjuntamente com a série do log da  $PTF_t$ .



**Figura 2:** Evolução da função de transição  $F(\cdot)$  (área),  $G(\cdot)$  (linha tracejada) e do log da  $PTF_{t-4}$  (linha cheia) — Estados Unidos

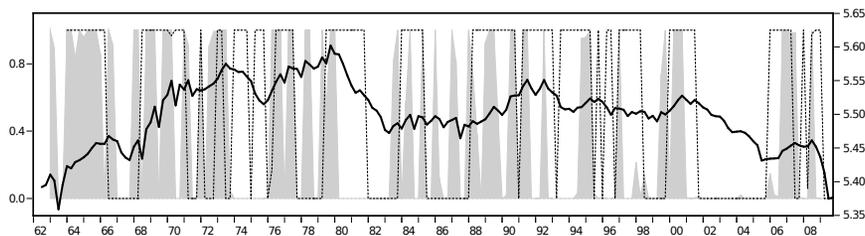


**Figura 3:** Evolução da função de transição  $F(\cdot)$  (área),  $G(\cdot)$  (linha tracejada) e do log da  $PTF_{t-4}$  (linha cheia) — Canadá



**Figura 4:** Evolução da função de transição  $F(\cdot)$  (área),  $G(\cdot)$  (linha tracejada) e do log da  $PTF_{t-4}$  (linha cheia) — Japão

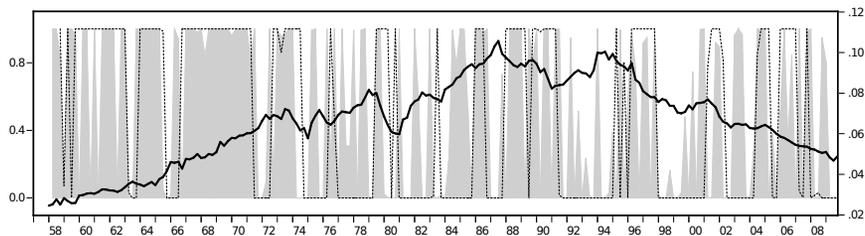
São perceptíveis os efeitos das exportações nos momentos em que a produtividade total dos fatores dessas economias encontra-se em trajetória de queda, observado quando  $F(\cdot)$  e  $G(\cdot)$  são inferiores a 0,5 nível em que ocorre a mudança de regime. Nos extremos das funções, essa dinâmica refere-se ao regime 1. De maneira geral, há reversão da queda do log da  $PTF_t$  nos momentos



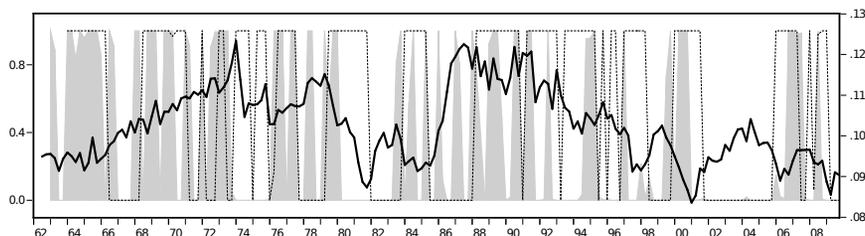
**Figura 5:** Evolução da função de transição  $F(\cdot)$  (área),  $G(\cdot)$  (linha tracejada) e do log da  $PTF_{t-2}$  (linha cheia) — Alemanha

em que  $F(\cdot)$  é inferior a 0,5 e  $G(\cdot)$  superior a 0,5 (regime 3). No caso do Canadá e Alemanha, verifica-se a situação inversa, uma restrição no crescimento do log da  $PTF_t$  quando  $F(\cdot)$  é superior a 0,5 e  $G(\cdot)$  inferior a 0,5 (regime 2).

O processo de ajustamento por meio das exportações só ocorre para o modelo MR-STVEC do Japão e Alemanha. Para salvar espaço, nas Figuras 6 e 7 serão destacados somente a evolução do  $EX_t$ , conjuntamente com as funções de transição para esse dois países.



**Figura 6:** Evolução da função de transição  $F(\cdot)$  (área),  $G(\cdot)$  (linha tracejada) e do índice  $EX_{t-4}$  (linha cheia) – Japão



**Figura 7:** Evolução da função de transição  $F(\cdot)$  (área),  $G(\cdot)$  (linha tracejada) e do índice  $EX_{t-2}$  (linha cheia) – Alemanha

Observa-se que nesses casos, de maneira geral, há movimento de ascensão de  $EX_t$  quando  $F(\cdot)$  é superior a 0.5 e  $G(\cdot)$  inferior a 0.5 (regime 2). Logo, como previsto no modelo estimado, a participação das exportações responde positivamente ao aumento da produtividade do país.

## 6 Conclusão

Este artigo constatou a existência de relação de longo prazo, nos últimos 50 anos, entre produtividade total dos fatores e exportações, para quatro países

de renda per capita elevada, mas com distintos graus de participação do comércio internacional no PIB — Estados Unidos, Canadá, Japão e Alemanha. Diferentemente de grande parte da literatura econômica, optou-se por avaliar as exportações através de sua participação no valor total exportado pelo mundo. A partir do teste de especificação proposto, composto por dois procedimentos, verificou-se que a dinâmica de ajustamento pode ser caracterizada por meio do modelo MR-STVEC. Essa última especificação tem a vantagem sobre o modelo linear por tratar isoladamente os movimentos de desequilíbrio, seja dos desvios associados à incidência de choques negativos e positivos de produtividade, tal como dos retornos à trajetória de longo prazo.

No caso dos Estados Unidos, uma economia com baixa participação do setor exportador no PIB, o modelo linear indica que, independente do tipo de choque, a PTF sempre ajusta. Logo, as exportações agem de modo a restringir potenciais choques positivos de produtividade, o que é pouco factível para uma economia com mercado doméstico amplo. Por outro lado, o modelo não linear evidencia um processo condicionado de orientação da produtividade via exportações. Ou seja, a PTF ajusta somente quando há choques negativos de produtividade. Já o caso canadense também inclui dinâmica de ajustamento via PTF quando há choques positivos. Portanto, as exportações agem tanto de modo a sustentar a produtividade tal como restringi-la, com diferenças somente no que tange à velocidade do ajuste. Esse resultado é compatível para uma economia pequena como o Canadá que possui maior influência do setor exportador na economia.

O sentido de causalidade reversa que se processa da PTF para as exportações vem a ocorrer somente para Japão e Alemanha. O Japão, semelhantemente aos Estados Unidos, é uma das maiores economias do mundo, possui baixa participação das exportações em seu produto e tal como evidenciado pelo modelo MR-STVEC estimado, a orientação da PTF via exportações só ocorre quando há choques negativos de produtividade. Por outro lado, choques positivos de produtividade tem efeito sobre as exportações, de modo a elevar sua participação no total exportado no mundo.

O modelo estimado para Alemanha, tal como para os demais países, possui o mecanismo de sustentação da produtividade quando há choques negativos. Entretanto, com a incidência de choques positivos, tanto a PTF como as exportações ajustam, sendo que a velocidade de ajustamento da primeira é superior. Essa dinâmica de causalidade pode ser explicada, pois tal como Japão e Estados Unidos, a Alemanha é uma das maiores economias do mundo. Contudo, possui participação maior do setor exportador no produto, similar ao caso canadense. Observamos, portanto, que esse país não possui mecanismo exclusivo de ajustamento da PTF orientado pelas exportações. Inovações no setor que não atua no comércio internacional podem também influenciar o setor exportador, promovendo padrão de causalidade reversa.

## **Agradecimentos**

Os autores agradecem os comentários e sugestões dos pareceristas da Revista Economia Aplicada, e dos professores Rodrigo de Losso da Silveira Bueno, Carlos Eduardo Soares Gonçalves, Sérgio Kannebley Júnior e Jaylson Jair da Silveira.

## Referências Bibliográficas

- Anderson, H. M. & Teräsvirta, T. (1992), 'Characterizing nonlinearities in business cycles using smooth transition autoregressive models', *Journal of Applied Econometrics* 7(S1), S119–S136.
- Aw, B. Y., Chen, X. & Roberts, M. J. (1997), 'Firm-level evidence on productivity differentials, turnover, and exports in taiwanese manufacturing', NBER Working Paper n. 6235.
- Awokuse, T. O. (2007), 'Causality between exports, imports, and economic growth: Evidence from transition economies', *Economics Letters* 94(3), 389–395.
- Balassa, B. (1978), 'Exports and economic growth: Further evidence', *Journal of Development Economics* 5(2), 181–189.
- Bernard, A. B. & Jensen, J. B. (1999), 'Exporting and productivity', National Bureau of Economic Research. Working Paper n. 7135.
- Bernard, A. B., Jensen, J. B., Redding, S. J. & Schott, P. K. (2007), 'Firms in international trade', *The Journal of Economic Perspectives* 21(3), 105–130.
- Bhagwati, J. N. (1998), *Protectionism*, 1 edn, MIT Press.
- Bruton, H. J. (1989), Import substitution, in H. CHENERY & T. SRINIVASAN, eds, 'Handbook of Development Economics', Vol. 2, Oxford:North-Holland.
- Clerides, S., Lach, S. & Tybout, J. (1998), 'Is learning by exporting important? micro-dynamic evidence from colombia, mexico, and morocco', *Quarterly Journal of Economics* 113(3), 903–947.
- Coe, D. & Helpman, E. (1995), 'International r&d spillovers', *European Economic Review* 39(5), 859–887.
- Cooley, T. & Prescott, E. (1995), Economic growth and business cycles, in T. COOLEY & E. PRESCOTT, eds, 'Frontiers of Business Cycle Research', Vol. 1, Princeton University Press.
- De Loecker, J. (2007), 'Do exports generate higher productivity? evidence from slovenia', *Journal of International Economics* 73(1), 69–98.
- Dijk, D. V. & Franses, P. H. (1999), 'Modeling multiple regime in the business cycle', *Macroeconomic Dynamics* 3(3), 311–340.
- Dijk, D. V., Teräsvirta, T. & Franses, P. H. (2002), 'Smooth transition autoregressive models — a survey of recent developments', *Econometric Reviews* 21(1), 1–47.
- Edwards, S. (1993), 'Openness, trade liberalization, and growth in developing countries', *Journal of Economic Literature* 31(3), 1358–1393.
- Edwards, S. (1998), 'Openness, productivity and growth: what do we really know?', *Economic Journal* 108(447), 383–398.
- Eitrheim, O. & Teräsvirta, T. (1996), 'Testing the adequacy of smooth transition autoregressive models', *Journal of Econometrics* 74(1), 59–75.

- Esfahani, H. S. (1991), 'Exports, imports, and economic growth in semi-industrialized countries', *Journal of Development Economics* 35(1), 93–116.
- Feder, G. (1982), 'On exports and economic growth', *Journal of Development Economics* 12(1), 59–73.
- Giles, J. A. & Williams, C. L. (2000a), 'Export-led growth: a survey of the empirical literature and some non-causality results. part 1', *Journal of International Trade & Economic Development* 9(4), 61–337.
- Giles, J. A. & Williams, C. L. (2000b), 'Export-led growth: a survey of the empirical literature and some non-causality results. part 2', *Journal of International Trade & Economic Development* 9(4), 445–470.
- Granger, C. W. J. & Teräsvirta, T. (1993), *Modeling Nonlinear Economic Relationships*, 1 edn, New York: Oxford University Press.
- Greenaway, D. & Sapsford, D. (1994), 'Exports, growth, and liberalization: An evaluation', *Journal of Policy Modeling* 16(2), 165–186.
- Grossman, G. M. & Helpman, E. (1991), *Innovation and Growth in the Global Economy*, 1 edn, Cambridge: MIT Press.
- Hall, R. E. & Jones, C. I. (1999), 'Why do some countries produce so much more output per worker than others?', *The Quarterly Journal of Economics* 114(1), 83–116.
- Henriques, I. & Sadorsky, P. (1996), 'Export-led growth or growth-driven exports? the canadian case', *The Canadian Journal of Economics* 29(3), 540–555.
- Issler, J. V., Pessoa, S. A. & Ferreira, P. C. (2004), 'Testing production functions used in empirical growth studies', *Economic Letters* 83(1), 29–35.
- Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of Economics Dynamics and Control* 12(2-3), 231–365.
- Johansen, S. (1991), 'Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models', *Econometrica* 59(6), 1551–1580.
- Krueger, A. (1980), 'Trade policy as an input development', *American Economic Review* 70(2), 288–292.
- Krugman, P. (1984), Import protection as export promotion, in H. Kierzkowski, ed., 'Monopolistic Competition in International Trade', Vol. 1, Oxford University Press.
- Kunst, R. M. & Marin, D. (1989), 'On exports and productivity: A causal analysis', *The Review of Economics and Statistics* 71(4), 699–703.
- Lundbergh, S., Dijk, D. V. & Teräsvirta, T. (2003), 'Time-varying smooth transition autoregressive models', *Journal of Business & Economic Statistics* 21(1), 104–121.
- Mackinnon, J. G., Alfred, A. H. & Michelis, L. (1999), 'Numerical distribution functions of likelihood ratio tests for cointegration', *Journal of Applied Econometrics* 14(5), 563–577.

Marin, D. (1992), 'Is the export-led growth hypothesis valid for industrialized countries?', *The Review of Economics and Statistics* 74(4), 678–688.

Rodrik, D. (1995), 'Trade strategy, investment and export: Another look at east asia', National Bureau of Economic Research Working Paper 5339.

Rothman, P., Dijk, V. D. & Franses, P. H. (2001), 'Multivariate star analysis of money-output relationship', *Macroeconomics Dynamics* 5(4), 506–532.

Teräsvirta, T. (1994), 'Specification, estimation, and evaluation of smooth transition autorregressive models', *Journal of the American Statistical Association* 89(425), 208–218.

Vamvakidis, A. (2002), 'How robust is the growth-openness connection? a historical evidence', *Journal of Economic Growth* 7(1), 57–80.

Xu, Z. (1996), 'On the causality between export growth and gdp growth: An empirical reinvestigation', *Review of International Economics* 4(2), 172–184.

Yamada, H. (1988), 'A note on the causality between export and productivity: an empirical re-examination', *Economics Letters* 61(1), 111–114.

## Apêndice A

Considerando o modelo em (10), o cálculo de  $\nabla H(z_{t-1}, \Delta x_{t-j}, s_{F_{t-d}}; \hat{\Psi}) = \frac{\partial H(z_{t-1}, \Delta x_{t-j}, s_{F_{t-d}}; \hat{\Psi})}{\partial \hat{\Psi}}$  é dado por:

$$\frac{\partial H}{\partial \hat{\mu}_1} = 1 \quad (10)$$

$$\frac{\partial H}{\partial \hat{\alpha}_1} = z_{t-1} \quad (11)$$

$$\frac{\partial H}{\partial \hat{\Phi}_{1,j}} = \Delta x_{t-j} \quad (12)$$

$$\frac{\partial H}{\partial \hat{\mu}_2} = F(s_{F_{t-d}}; \hat{\gamma}_F, \hat{c}_F) \quad (13)$$

$$\frac{\partial H}{\partial \hat{\alpha}_2} = z_{t-1} F(s_{F_{t-d}}; \hat{\gamma}_F, \hat{c}_F) \quad (14)$$

$$\frac{\partial H}{\partial \hat{\Phi}_{2,j}} = \Delta x_{t-j} F(s_{F_{t-d}}; \hat{\gamma}_F, \hat{c}_F) \quad (15)$$

$$\begin{aligned} \frac{\partial H}{\partial \hat{\gamma}_F} = & \left(1 + \exp\{-\hat{\gamma}_F (s_{F_{t-d}} - \hat{c}_F)\}\right)^{-2} \exp\{-\hat{\gamma}_F (s_{F_{t-d}} - \hat{c}_F)\} * \\ & \left(s_{F_{t-d}} - \hat{c}_F\right) \left(\hat{\mu}_2 + \hat{\alpha}_2 z_{t-1} + \sum_{j=1}^p \hat{\Phi}_{2,j} \Delta x_{t-j}\right) \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{\partial H}{\partial \hat{c}_F} = & \hat{\gamma}_F \left(1 + \exp\{-\hat{\gamma}_F (s_{F_{t-d}} - \hat{c}_F)\}\right)^{-2} \exp\{-\hat{\gamma}_F (s_{F_{t-d}} - \hat{c}_F)\} * \\ & \left(\hat{\mu}_2 + \hat{\alpha}_2 z_{t-1} + \sum_{j=1}^p \hat{\Phi}_{2,j} \Delta x_{t-j}\right) \end{aligned} \quad (17)$$

## Apêndice B

**Tabela B.1:** Teste Traço de Razão de Verossimilhança

País	n	p	Estatística Traço	Vetor de Cointegração
Estados Unidos	212	4	16,3**	[1 – 1,46 – 6,14]
Canadá	212	4	15,1*	[1 – 7,56 – 5,96]
Japão	212	4	20,7**	[1 – 3,13 – 8,79]
Alemanha	192	4	20,5**	[1 – 5,73 – 4,91]

Notas: Os valores críticos teste traço são obtidos de Mackinnon et al. (1999). \*  
significante a 10%, \*\* significante a 5% e \*\*\* significante a 1%. Os modelos estimados  
para Estados Unidos e Canadá incluem constante no vetor de correção de erros.



# IMPACTOS DA ALTERAÇÃO DE PERFIL DO CONSUMO DAS FAMÍLIAS SOBRE O EMPREGO POR OCUPAÇÃO E NÍVEL DE QUALIFICAÇÃO: APLICAÇÃO DE UM MODELO MULTISSETORIAL PARA O BRASIL

ADELAR FOCHEZATTO \*  
CARLOS EDUARDO LOBO E SILVA †

## Resumo

Este trabalho analisa os efeitos da alteração do perfil do consumo das famílias, verificada nos últimos anos sobre o mercado de trabalho. Mais especificamente, tem como objetivo: a) verificar as mudanças ocorridas na estrutura da demanda final; e b) dimensionar os efeitos das mudanças na estrutura da demanda final sobre o emprego por atividade econômica, ocupação e qualificação. Para atingir os objetivos, utiliza-se um modelo multissetorial. Os resultados mostram que o crescimento e alteração do padrão dos gastos das famílias impulsionam especialmente a demanda pelo trabalho não qualificado.

**Palavras-chave:** Gastos das famílias; Demanda por mão de obra; Qualificação da mão de obra; Ocupações.

## Abstract

This paper analyzes the effects on the labor market coming from the change of consumption pattern in Brazil observed during the last decade. More specifically, the paper has two objectives: a) identifying the structural changes of the final demand; and b) measuring the effects of the final-demand structural changes on the number of jobs by economic sectors, occupation, and qualification. To reach those objectives, a multi-sectoral model was used in this article. The main result shows the final-demand structural changes boost the demand for unskilled workers.

**Keywords:** Consumption; Labor demand; Labor qualification; Job occupation.

**JEL classification:** D57, J23, J24, O15.

**DOI:** <http://dx.doi.org/10.1590/1413-8050/ea122499>

---

\* Doutor em Economia. Professor Titular da PUCRS. Pesquisador do CNPq. E-mail: [adelar@pucrs.br](mailto:adelar@pucrs.br)

† PhD em Planejamento Regional. Professor da PUCRS. Pesquisador do CNPq. E-mail: [carlos.silva@pucrs.br](mailto:carlos.silva@pucrs.br)

## 1 Introdução

A economia brasileira tem apresentado níveis de crescimento relativamente elevados na última década em comparação com as duas décadas anteriores. Conforme dados das Contas Nacionais do IBGE, enquanto no período de 1980 a 2002 a taxa geométrica média de crescimento real do Produto Interno Bruto foi de 2,0% ao ano, no período de 2002 a 2009 a taxa média de crescimento foi de 3,5% ao ano. O ritmo de crescimento econômico da primeira década do século contou, entre outros fatores, com políticas que favorecem direta ou indiretamente a expansão do consumo das famílias. Bielschowsky (2012) as divide em quatro grupos: rápido aumento da massa salarial; intensificação nas transferências de renda aos mais pobres; estabilidade de preços de bens industriais populares; e forte ampliação do crédito. Adicionalmente, as desonerações setoriais promovidas pelo governo federal e o aumento contínuo da participação da mulher no mercado de trabalho (Jacinto & Caetano 2011) nos ajudam a entender a expansão do consumo das famílias.

Muitos desses fatores afetam de maneira bastante desigual as diferentes classes de renda da sociedade brasileira. A expansão do crédito,<sup>1</sup> aliado a um expressivo aumento do acesso à rede bancária (Barone & Sader 2008, Silva & Moraes 2013), e o crescimento da transferência de renda — seja por meio de programas como Bolsa Família, seja por impacto do aumento do salário mínimo real sobre as pensões de aposentadoria — são fatores que sugerem uma alteração significativa no padrão de consumo das famílias brasileiras.

No entanto, o objetivo deste artigo não se encerra na análise do padrão de consumo; a motivação do trabalho tem nas alterações de consumo das famílias seu ponto de partida. Retomando uma preocupação apresentada por Bresser-Pereira (2004) sobre um possível ciclo de desenvolvimento concentrador de renda, que vincula o padrão de consumo com a demanda por mão de obra, nosso trabalho ambiciona estimar o impacto que as alterações recentes do padrão de consumo das famílias brasileiras provocam na demanda por mão de obra, classificada aqui quanto às ocupações e à qualificação. Mais especificamente, objetiva-se: a) verificar as mudanças ocorridas na estrutura da demanda final no Brasil e b) dimensionar os efeitos das mudanças na estrutura da demanda final sobre o emprego por atividade econômica, por ocupação e por qualificação. Para atingir o objetivo proposto, é utilizado um modelo multissetorial.

Os dados utilizados para aplicar o modelo são os da nova Matriz de Insumo Produto (MIP) — 2005 e das contas nacionais; os dados da Pesquisa de Orçamentos Familiares (POF) de 2002-2003 e de 2008-2009, ambas do IBGE; além da Relação Anual de Informações Sociais (RAIS) de 2008, disponíveis no Ministério do Trabalho e Emprego. Com os dados da MIP e das contas nacionais é elaborada uma matriz de contabilidade social atualizada para o Brasil e, com os dados da POF, elabora-se uma matriz de consumo setorial das famílias por classe de renda. Os dados da RAIS são usados para desagregar o emprego por ocupações e por qualificação. Para a aplicação do modelo, faz-se ainda a compatibilização setorial das informações entre os setores da RAIS, da POF e do Sistema de Contas Nacionais.

---

<sup>1</sup>As operações de crédito em relação ao PIB brasileiro passaram de 24,2% em março de 2004 para 48,5% em outubro de 2011 citepGaleano2011, com impulso especialmente significativo no segmento pessoas físicas, que saltou de um total correspondente a 5,6% do PIB em 2004 para 14,9% no final de 2009 (Torres 2010).

Não parece haver razão para acreditar em um retrocesso quanto às políticas de apoio ao fortalecimento do mercado interno, especialmente em relação àquelas que têm efeitos diversos para diferentes classes de renda. Portanto, entender como e em que magnitude esse novo perfil de consumo impacta o mercado de trabalho constitui um passo relevante para a compreensão mais ampla sobre tendências, oportunidades e dificuldades recentes da economia brasileira. Além disso, os resultados obtidos com o método usado neste trabalho podem ajudar a identificar alguns desafios para a qualificação da força de trabalho no Brasil e auxiliar no desenho de políticas que possam promover um crescimento econômico com melhoria da equidade e com aumento da competitividade.

Os estudos que usam a metodologia adotada neste trabalho em geral procuram estabelecer relações estruturais entre a composição da demanda final, a estrutura setorial da produção, o perfil de remunerações dos fatores de produção e a distribuição da renda entre grupos de famílias. Essas dimensões econômicas são intermediadas por um conjunto de relações formando um fluxo circular da renda. Nesse fluxo, não há um ponto inicial e cada uma das suas partes interfere sobre as demais. Os trabalhos pioneiros são Pyatt & Round (1977, 1979), Stone (1985), Pyatt (1988), Thorbecke (1995), Pyatt (2001) entre outros. Em geral, a hipótese central desses trabalhos é de que o perfil da distribuição da renda, medidos pelo padrão de consumo, tem diferentes efeitos sobre os setores produtivos, influenciando a estrutura produtiva da economia.

Cruyce & Wera (2007) utilizam uma matriz de insumo-produto da Bélgica de 2000 e 2002 para estimar os multiplicadores de emprego por nível de educação, idade, sexo e por atividade econômica. Os autores concluem que existe uma diferença significativa entre os setores em termos de impactos diretos, indiretos e induzidos sobre o trabalho, desagregado pelas características citadas acima. Por exemplo, através dos efeitos indiretos e induzidos, um setor que demanda emprego de baixa qualificação pode ocasionar aumentos importantes na demanda de emprego qualificado.

Entre os trabalhos aplicados para a economia brasileira destacam-se Bonelli & Cunha (1981, 1982, 1983), os quais, em três artigos diferentes, utilizaram modelos alternativos de insumo-produto para analisar a reação da produção setorial a diferentes hipóteses de distribuição de renda e do consumo entre 1970 e 1975.

Entre os estudos mais recentes, Guilhoto et al. (1996) analisaram as estruturas de produção, consumo e distribuição de renda no Brasil em 1975 e 1980, através das matrizes ajustadas de insumo-produto para esses anos. Seus resultados indicam que o polo dinâmico da economia se encontra nos consumidores da classe de renda entre 5 e 20 salários mínimos. Há, também, o trabalho de Cavalcanti (1997), que analisou os impactos de mudanças na distribuição setorial da renda sobre a distribuição funcional da renda. Seus resultados indicam que a estrutura produtiva em vigor propicia uma canalização da renda adicional em favor da classe dos rentistas em prejuízo dos assalariados. Ele conclui, também, que os resultados são muito diferentes entre os setores em função das suas ligações intersetoriais, dos seus coeficientes de consumo e de renda. Por fim, Fochezatto (2011), utilizando dados da POF 2002-03, analisa os efeitos de diferentes estruturas de demanda final sobre a distribuição de renda entre grupos de famílias. Seus resultados mostram que um aumento relativo do consumo das famílias mais pobres não garante uma melhoria na distribuição de renda.

A proposta deste trabalho segue a linha metodológica dos trabalhos acima, especialmente Cruyce & Wera (2007) e Fochezatto (2011), embora com uma perspectiva temática um pouco diferente. A pergunta central do presente artigo é a seguinte: mudanças no padrão do consumo e, por consequência, na produção dos setores, influenciam o emprego em quais setores, ocupações e níveis de qualificação? O marco lógico tem as seguintes etapas: políticas governamentais e dinâmicas socioeconômicas alteram a estrutura da demanda final, a qual afeta diferentemente a produção dos setores. Os setores possuem diferentes tecnologias de produção e, portanto, geram diferentes impactos sobre o emprego, ocupações e qualificação.

## 2 Metodologia

A apresentação da metodologia está dividida em duas partes. Inicialmente, o modelo proposto é detalhado, vinculando os passos da sua construção aos objetivos do trabalho. Em seguida, apresenta-se o tratamento dos dados de forma mais detalhada, identificando as estratégias adotadas e hipóteses assumidas.

### 2.1 Construção do Modelo

As alterações na demanda, no perfil produtivo e na estrutura de rendimentos da economia influenciam vários mercados simultaneamente. Para analisar seus efeitos, os modelos multissetoriais são instrumentos adequados porque eles capturam as principais interações entre os diferentes mercados existentes no sistema econômico.

A matriz de insumo-produto é uma representação estática da estrutura econômica de uma economia em um determinado período de tempo. Ela mostra as relações intersetoriais, a demanda final dos setores, o valor adicionado das atividades produtivas e os impostos indiretos, capturando, assim, a interdependência entre os mercados de fatores e de produtos. As transações são organizadas de maneira consistente, de modo a igualar receitas e despesas para cada agente econômico do sistema. As linhas da matriz mostram a demanda total do produto de cada setor, a qual é igual à soma da demanda intermediária e da demanda final (equação 1). As colunas mostram que a produção bruta de cada setor é igual ao consumo intermediário mais os fatores primários ou valor adicionado (equação 2). Para cada setor da economia, se cumpre que a sua produção bruta é igual à sua demanda total (equação 3). Em termos agregados, a soma do valor adicionado dos setores fornece a renda agregada da economia e a soma da demanda final dos setores resulta no dispendio agregado (equação 4). Em termos formais:

$$X_i = \sum_{j=1}^n X_{ij} + \sum_{s=1}^S Y_{is} \quad (1)$$

$$X_j = \sum_{i=1}^n X_{ij} + \sum_{r=1}^R V_{rj} \quad (2)$$

$$\sum_{j=1}^n X_{ij} + \sum_{s=1}^S Y_{is} = \sum_{i=1}^n X_{ij} + \sum_{r=1}^R V_{rj} \quad (3)$$

$$\sum_{i=1}^n \sum_{s=1}^S Y_{is} = \sum_{j=1}^n \sum_{r=1}^R V_{rj} \quad (4)$$

onde:  $X_i$  é a demanda total do produto do setor  $i$ ;  $X_{ij}$  é a demanda intermediária do produto do setor  $i$  por parte do setor  $j$ ;  $Y_{is}$  é a demanda final do produto do setor  $i$  na categoria  $s$ ;  $X_j$  é a produção bruta do setor  $j$ ;  $V_{rj}$  é o fator primário  $r$  no valor adicionado do setor  $j$ ;  $i, j = 1, 2, \dots, n$  são os setores produtivos da economia;  $s = 1, 2, \dots, S$  são as categorias de demanda final; e  $r = 1, 2, \dots, R$  são os fatores primários que compõem o valor adicionado.

No modelo de insumo-produto, supõe-se que os coeficientes técnicos de produção são fixos e mostram a quantidade do produto do setor  $i$  requerida para produzir uma unidade de produto do setor  $j$ . Os requerimentos de fatores primários também têm uma relação fixa e definem a quantidade do fator primário  $r$  requerida para produzir uma unidade de produto total no setor  $j$ . Assim:

$$a_{ij} = \frac{X_{ij}}{X_j} \text{ ou } a_{ij}X_j = X_{ij} \quad (5)$$

$$b_{rj} = \frac{V_{rj}}{X_j} \text{ ou } b_{rj}X_j = V_{rj} \quad (6)$$

onde:  $a_{ij}$  representa os coeficientes técnicos de produção; e  $b_{rj}$  representa os coeficientes dos fatores primários.

No modelo aberto de insumo-produto, todas as categorias de demanda final são consideradas exógenas. Partindo das equações (1) e (5), e usando a notação matricial, a solução desse modelo é definida pela seguinte expressão:

$$X = (I - A)^{-1} Y \quad (7)$$

onde:  $X$  é um vetor coluna ( $nx1$ ) de produção total dos setores;  $I$  é uma matriz identidade ( $nxn$ );  $A$  é a matriz de coeficientes técnicos ( $nxn$ ); e  $Y$  é um vetor coluna ( $nx1$ ), representando o somatório das categorias de demanda final para o produto de cada setor. Os coeficientes da matriz inversa  $(I - A)^{-1}$  são chamados de requerimentos diretos e indiretos de produção.

Além dos requerimentos de produção, é possível definir os requerimentos de fatores primários. Combinando a equação (6) na forma matricial com a equação (7) resulta na seguinte expressão de demanda de fatores primários:

$$V = BX \text{ ou } V = B(I - A)^{-1} Y \quad (8)$$

onde:  $V$  é um vetor indicando a demanda de um determinado fator primário em cada setor econômico ( $nx1$ );  $B$  é uma matriz diagonal dos coeficientes de um determinado fator primário ( $nxn$ ). A matriz  $B(I - A)^{-1}$  é chamada de matriz de requerimentos diretos e indiretos de fatores primários.

Neste trabalho, foi utilizado um modelo de insumo-produto fechado em que o consumo das famílias foi considerado endógeno.<sup>2</sup> Substituindo a equação (5) na equação (1), esta equação, na forma matricial, pode ser escrita conforme a equação (9). Esta, por sua vez, com a separação do consumo das

<sup>2</sup>Essa transformação do modelo aberto para o fechado baseia-se em Najberg & Vieira (1996) e Najberg & Ikeda (1999). Vale lembrar que, nesse caso, a alteração da demanda tem impacto direto, indireto e pelo efeito renda.

famílias das demais categorias de demanda final, pode ser reescrita de acordo com a equação (10):

$$AX + Y = X \quad (9)$$

$$AX + C + Y' = X \quad (10)$$

em que  $Y$  é um vetor da soma dos componentes da demanda final ( $nx1$ );  $C$  é o vetor de consumo das famílias ( $nx1$ ); e  $Y'$  é um vetor da soma das demais categorias de demanda final ( $nx1$ ).

Considerando-se que o consumo do produto de cada setor da economia seja endógeno e uma função linear da renda, então ele pode ser expresso conforme a equação (11). Sabendo-se que a renda agregada (valor adicionado) é igual ao dispêndio agregado (equação 4) e que a renda setorial representa uma fração do valor bruto da produção do setor, esta equação pode ser reescrita de acordo com a equação (12). A equação (13) nada mais é do que a equação (12) escrita no formato matricial:

$$X_i = d_i Y \quad (11)$$

$$C_i = d_i \sum_{j=1}^n r_j X_j \quad (12)$$

$$C = DRX \quad (13)$$

onde  $C_i$  é o consumo do produto do setor  $i$ ;  $d_i$  é um vetor de propensão média a consumir do produto  $i$ ;  $r_j$  é o coeficiente de valor adicionado do setor  $j$ .

Substituindo a equação (13) na equação (10) pode-se definir a solução do modelo fechado para a produção dos setores (equação 14) e para a demanda de fatores primários (equação 15):

$$X = (I - A - DR)^{-1} Y1 \quad (14)$$

$$V = B(I - A - DR)^{-1} Y' \quad (15)$$

Este trabalho calculou os impactos sobre o emprego usando o modelo da equação (16). As variações na demanda final em cada setor da matriz foram definidas pela diferença entre a Pesquisa de Orçamento Familiar (POF) do ano 2008-09 e da POF de 2002-03. A partir dos resultados sobre o emprego total dos setores, foi feita a desagregação para emprego por qualificação (grau de instrução) e por ocupações, usando as equações (17) e (18).

$$\frac{\Delta V}{\Delta Y'} = B(I - A - DR)^{-1} \quad (16)$$

$$V_{qj} = h_{qj} \Delta V_j \quad (17)$$

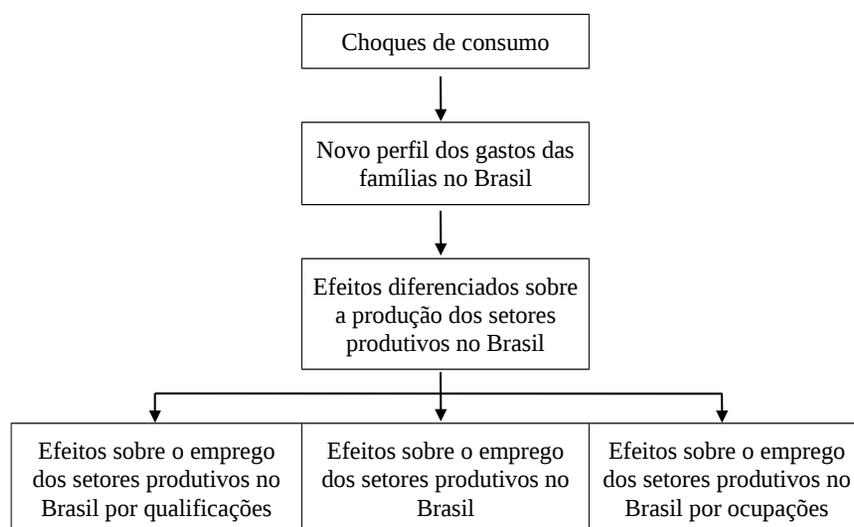
$$V_{oj} = h_{oj} \Delta V_j \quad (18)$$

onde:  $\Delta V$  é o impacto sobre o emprego do setor  $j$ , decorrente de uma alteração na demanda final  $\Delta Y'$ ;  $V_{qj}$  é a variação do emprego da qualificação  $q$  no setor  $j$ ;

$h_{qj}$  é a participação do emprego da qualificação  $q$  no setor  $j$ ;  $\Delta V_j$  é a variação do emprego total no setor  $j$ ;  $V_{oj}$  é a variação do emprego da ocupação  $o$  no setor  $j$ ; e  $h_{oj}$  é a participação do emprego da ocupação  $o$  no setor  $j$ .

Em linhas gerais, a mudança de perfil dos gastos das famílias será captada pela Pesquisa de Orçamento Familiar dos anos 2002-03 e 2008-09. Essa variação entra como um choque na matriz de insumo-produto que, a partir dos seus multiplicadores setoriais de produção e dos índices também setoriais do número de trabalhadores gerado por unidade monetária produzida, fornece o impacto na demanda por mão de obra. Finalmente, a RAIS informa as ocupações e qualificações do trabalho de cada setor produtivo. Portanto, a metodologia parte da identificação da mudança de perfil dos gastos das famílias, estima os efeitos sobre a produção e, finalmente, estima como a alteração na produção afeta a demanda por trabalho em termos de ocupações e de qualificações (mensurada neste trabalho pelo nível de escolaridade).

A Figura 1 mostra de forma esquematizada a estratégia empírica que será seguida na elaboração da pesquisa.



Fonte: Elaboração dos autores.

**Figura 1:** Fluxograma com os procedimentos operacionais da execução da pesquisa

A utilização de diferentes fontes de dados exigiu um esforço de compatibilização e ajustes que são descritos na seção abaixo.

## 2.2 O tratamento dos dados

Os dados utilizados da POF foram relativos aos períodos 2002-2003 e 2008-2009. Os valores do POF de 2002-03 foram corrigidos pelo IPCA, considerando a variação do nível geral de preços de dezembro de 2002 a dezembro de 2008. Em seguida, calculou-se a variação dos gastos entre os dois períodos.

A próxima tarefa foi compatibilizar os produtos de consumo da POF com os setores da matriz insumo-produto de 2005 com 55 produtos, que foi feita a partir da análise da descrição dos setores. Três aspectos merecem menção.

Os gastos com alimentação da POF foram examinados em detalhes a partir da tabela que traz a distribuição das despesas mensal familiar com alimentação. Assim, com base na maior desagregação apresentada na tabela da POF, o montante total de alimentação foi distribuído entre os três setores da matriz — agricultura (12,7%), alimentos e bebidas (67,6%) e serviços de alimentação e alojamento (19,7%).

Vale dizer ainda que em alguns casos o consumo da POF incluía mais de um setor. Especialmente na linha “outros” de cada área. Por exemplo, Serviços e Taxas incluem como subprodutos energia elétrica, gás, água, esgoto, telefone fixo, telefone celular, pacote de telefone, TV e outros. Enquanto os quatro primeiros pertencem ao setor 0401 — Eletricidade e gás, água, esgoto e limpeza urbana; os demais ligados à telefonia e internet pertencem ao setor 0801 — Serviços de Informação. Assim, a linha “Outros” foi distribuída pelos dois setores com participações iguais.

Finalmente, um ponto relevante do procedimento adotado diz respeito aos valores das duas linhas relativas ao “Aumento do ativo”: imóvel (aquisição) e imóvel (reforma). A soma das variações de aquisição e reforma foi considerada como variação do setor de construção. Entretanto, considerando a matriz insumo-produto, nota-se que as famílias não consomem nada do setor de construção. Isso se dá pelo fato das contas nacionais considerarem os gastos das famílias em construção e reforma de imóvel como parte da formação bruta de capital, constituindo-se como parte do investimento. Dessa forma, apesar dos gastos das famílias na construção e reforma de imóvel não serem considerados consumo pelas contas nacionais, eles impactam diretamente o setor de construção. Por essa razão, os valores do item “Aumento do ativo” da POF foi puxado na matriz de insumo-produto para a coluna de demanda final das famílias e para a linha relativa ao setor de construção. Em outras palavras, consideram-se aqui os gastos relativos à construção e reforma de imóveis como demanda final das famílias pelo setor da construção.

Superados os cálculos e compatibilizações da POF, passou-se à construção dos coeficientes técnicos da matriz insumo-produto. O primeiro passo foi construir um modelo completo, que levasse em conta não apenas as relações intersetoriais, mas que também considerasse o valor adicionado da produção que impulsiona o consumo das famílias que, por sua vez, retorna como aumento da demanda. Para tanto, um vetor coluna de consumo das famílias foi construído a partir das contas nacionais e inserido na matriz (incluindo no vetor o valor do aumento de ativo na linha relativa ao setor de construção civil), além da linha de valor adicionado que também passou a fazer parte da matriz completa.

A partir do choque de demanda, a matriz dos coeficientes fornece a variação do valor bruto da produção. Dividindo o valor total da produção de 2005 pelo número de trabalhadores chega-se ao índice do valor bruto da produção setorial para cada trabalhador empregado no setor em questão. Como esses índices são utilizados para estimar variações de emprego em função da variação da demanda a preços de 2008-09, é necessário corrigir os valores, o que foi feito utilizando-se mais uma vez o IPCA.

O impacto da demanda no valor bruto da produção e os índices mencionados acima nos permitem identificar o número de postos de trabalho gerados / destruídos em função da variação da demanda final. De posse desses valores, a Relação Anual de Informações Sociais (RAIS) possibilita a decomposição do número de trabalhadores de cada setor em diferentes classes de ocupação e

qualificação, que é definida aqui pelo nível de escolaridade.

### **3 Resultados**

Os principais resultados apresentados neste artigo podem ser divididos em três grupos. Os resultados iniciais mostram o impacto da variação dos gastos das famílias entre 2002-03 e 2008-09 na demanda por mão de obra. Mais especificamente, os resultados trazem o impacto total desagregado em nove ocupações e em onze níveis de escolaridades, tratados, a partir daqui, como qualificações.

O segundo grupo de resultados recupera a estrutura da mão de obra empregada no Brasil em 2008, tanto em termos de ocupações quanto em qualificação, e a compara com a distribuição dos impactos causados pela variação dos gastos das famílias. Essa comparação enriquece a análise, pois permite identificar em quanto o aumento dos gastos das famílias durante o período estudado alterou a estrutura vigente e as características da mão de obra da economia brasileira.

Finalmente, o terceiro grupo de resultados analisa os efeitos da mudança do padrão de consumo na demanda por mão de obra. Para tanto, os resultados encontrados inicialmente são comparados com as alterações na demanda por mão de obra que ocorreriam, caso o padrão dos gastos das famílias de 2008-09 fosse o mesmo do padrão de 2002-03. Assim, considera-se a variação total ocorrida em termos absolutos, mas a distribui pelos itens de despesa conforme o padrão vigente em 2002-03. Esse exercício ressalta a relevância da mudança de perfil de gastos na demanda por mão de obra, não apenas em termos de números de postos de trabalho, mas também no tocante à sua distribuição em termos de ocupação e qualificação.

O impacto total no número de empregos em razão do aumento de consumo das famílias entre os períodos 2002-03 e 2008-09 foi de 15.866.069 postos de trabalho (última coluna dos anexos 1 e 2). Esse valor corresponde a 17,91% do total da mão de obra empregada em 2002.<sup>3</sup> A Tabela A.1 do Apêndice mostra a distribuição dos quase 16 milhões de empregos gerados nas nove ocupações definidas pela RAIS, e traz ainda o impacto por setor e por ocupações; enquanto, na Tabela A.2 do Apêndice, observa-se a distribuição por qualificação total e o impacto por setor e por qualificação.

Quanto às ocupações, destaque especial deve ser dado ao setor de trabalhadores da produção de bens e serviços industriais *I* (27,5%), seguido pelos trabalhadores do serviço do comércio (19,5%) e pelos trabalhadores agropecuários, florestais e da pesca (18,5%). Os trabalhadores do grupo da produção de bens e serviços industriais *I* “desempenham trabalhos ‘artesanais’, entendidos como os trabalhadores que conhecem e trabalham em todas as fases do processo de produção, independentemente de usar equipamentos rudimentares ou sofisticados” (Classificação Brasileira de Ocupações: CBO, 2010, p.5). Para diferenciá-los dos trabalhadores do grupo da produção de bens e serviços industriais *II*, estes últimos são os operadores de máquinas. Finalmente, os trabalhadores do grupo de serviços de reparação e manutenção são aqueles de nível elementar de competência (não qualificados) (CBO, 2010).

---

<sup>3</sup>Para manter a consistência dos resultados, o total de emprego deve vir da matriz insumo-produto. Como não há disponível a matriz para o ano de 2002, o total de empregos foi calculado a partir de uma média ponderada dos totais das matrizes de 2000 e 2005.

Em relação à qualificação, a variação dos gastos das famílias tem impactado de forma mais preponderante a demanda por trabalhadores com nível médio completo (34,3%). Em seguida, aparece o grupo fundamental completo (16,7%). Chama a atenção o pequeno impacto na demanda por trabalhadores com nível superior completo (6,2%) e a quase inexistente por trabalhadores com pós-graduação: o aumento da demanda pelas duas categorias — com mestrado e com doutorado — somadas não chegam a 0,2% do impacto total.

As Figuras 2 e 3 comparam a distribuição do impacto total com a estrutura de produção existente em 2008. Em ocupações (Figura 2), apesar do impacto percentual da demanda pelos trabalhadores agropecuários, florestais e da pesca ser apenas o terceiro mais significativo, como visto na Tabela A.1, quando comparado com a estrutura produtiva, a diferença passa a ser a mais significativa, pois somente 4% do total dos trabalhadores estão locados nestas atividades. A disparidade pode ser explicada pelo fato do item alimentação ter participação expressiva nas despesas totais das famílias. Mais especificamente, os gastos com alimentação representaram 16% da despesa total e 19,7% da despesa de consumo das famílias.

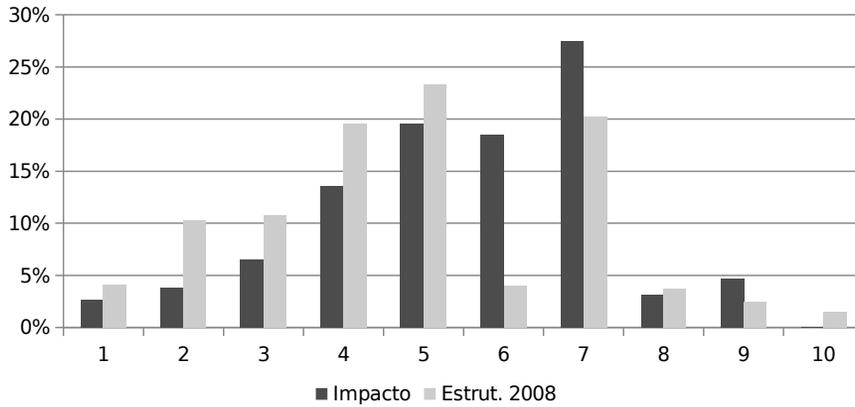
Vale mencionar ainda que as duas ocupações que são proporcionalmente mais demandadas do que as respectivas participações no total da mão de obra são os grupos relativos aos bens e serviços industriais *I* — aqueles que não operam máquinas — e serviços de manutenção e reparação, composto, segundo a própria definição da CBO (2010), por trabalhadores de baixa qualificação. Certamente, o importante aumento dos gastos na construção civil explica ao menos parte destes resultados. A participação da variação líquida de ativos via aquisição de imóvel em relação às despesas totais saltou de 2,77% em 2002-03 para 3,96% seis anos depois.

Já com relação às qualificações, as conclusões são bastante claras: não considerando os trabalhadores com pós-graduação — uma vez que tanto a participação desse grupo na estrutura produtiva quanto o impacto gerado sobre eles são pouco significativos — a distribuição do aumento de demanda está mais concentrada nos grupos de baixa qualificação do que a distribuição dos trabalhadores na estrutura produtiva. Dos nove grupos restantes, isso acontece nos seis grupos de menor qualificação e o contrário se verifica nos grupos de trabalhadores com médio completo, superior incompleto e superior completo.

Novas comparações são reportadas nas Figuras 4 e 45. O impacto calculado inicialmente em termos de qualificação e ocupação é agora comparado com o impacto que teria sido gerado, caso o padrão de despesa (distribuição entre os itens de despesa) verificado em 2002-03 se mantivesse até 2008-09. Assim, para calcular este efeito hipotético, distribui-se o a variação total observada entre 2002-03 e 2008-09, obedecendo à participação de cada item nos gastos totais de 2002-03. A diferença entre o impacto de demanda por trabalho inicialmente calculado e este hipotético, que “congela” o padrão dos gastos das famílias de 2002-03, é apresentada pelas Figuras 4 e 5. Na Figura 4, os valores são distribuídos por ocupação, enquanto a Figura 5 mostra a diferença em termos de qualificação da mão de obra.

Como primeiro resultado, chama a atenção o fato de que, tivesse o padrão de gastos das famílias se mantido ao longo do período estudado, a variação total de gastos ocorrida teria gerado 698.788 empregos a mais do que os 15.866.069 estimados neste trabalho, que considera a variação real de cada item da despesa.

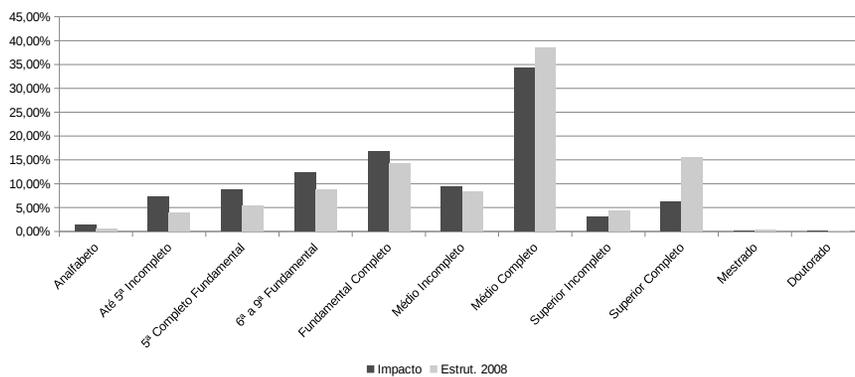
Analisando a Figura 4, percebe-se que o grupo ‘trabalhadores agropecuá-



Fonte: Resultados da pesquisa.

- <sup>1</sup> Membros superiores do poder público, dirigentes de organizações de interesse público.
- <sup>2</sup> Profissionais das ciências e das artes.
- <sup>3</sup> Técnicos de nível médio.
- <sup>4</sup> Trabalhadores de serviços administrativos.
- <sup>5</sup> Trabalhadores dos serviços, vendedores do comércio em lojas e mercados.
- <sup>6</sup> Trabalhadores agropecuários, florestais e da pesca.
- <sup>7</sup> Trabalhadores da produção de bens e serviços industriais I.
- <sup>8</sup> Trabalhadores da produção de bens e serviços industriais II.
- <sup>9</sup> Trabalhadores em serviços de reparação e manutenção.
- <sup>10</sup> Não classificado.

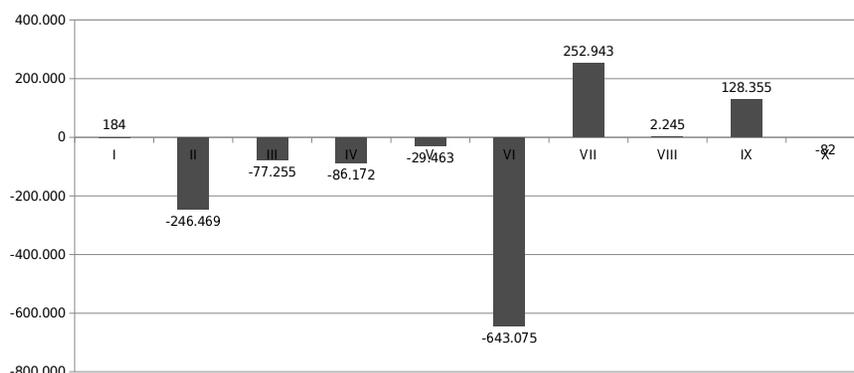
**Figura 2:** Impacto da variação dos gastos das famílias na demanda por trabalho comparado com a estrutura de produção — por ocupações.



Fonte: Resultados da pesquisa.

**Figura 3:** Impacto da variação dos gastos das famílias na demanda por trabalho comparado com a estrutura de produção — por qualificação.

rios, florestais e da pesca' apresenta a maior diferença, que ultrapassa os 600 mil empregos. O novo padrão de gasto das famílias demanda, proporcionalmente, menos trabalhadores desse grupo de trabalhadores do que demandava o padrão de 2002-03. Uma explicação para essa diferença parece ser a participação do item 'alimentos' nas despesas de consumo, que passou de 20,7% para 19,7%. À primeira vista, essa redução pode parecer pequena demais para explicar tal diferença, mas um ponto percentual na participação dos gastos totais de consumo altera significativamente os valores absolutos: enquanto a variação real dos gastos em alimentos apresentou ligeira redução entre 2002-03 e 2008-09, para manter os 20,7% de participação de 2002-03, ela teria que aumentar 4,7% durante o mesmo período. O padrão de gastos de 2008-09 gera maior demanda apenas para os trabalhadores de bens e serviços industriais I e reparo e manutenção (Na Figura 4, grupos VII e IX).



Fonte: Resultados da pesquisa.

I Membros superiores do poder público, dirigentes de organizações de interesse público.

II Profissionais das ciências e das artes.

III Técnicos de nível médio.

IV Trabalhadores de serviços administrativos.

V Trabalhadores dos serviços, vendedores do comércio em lojas e mercados.

VI Trabalhadores agropecuários, florestais e da pesca.

VII Trabalhadores da produção de bens e serviços industriais I.

VIII Trabalhadores da produção de bens e serviços industriais II.

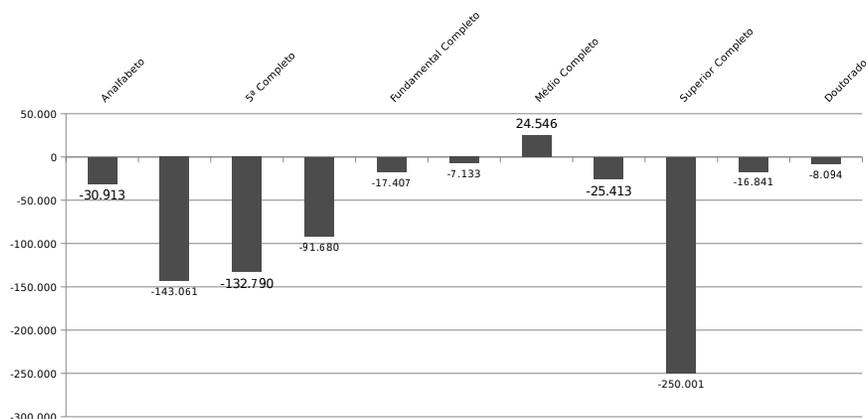
IX Trabalhadores em serviços de reparação e manutenção.

X Não classificado.

**Figura 4:** Impactos comparados: padrão de gastos das famílias de 2008 menos padrão de gastos das famílias de 2002 — por ocupação.

Finalmente, a Figura 5 mostra a mesma comparação para as qualificações. Com exceção do grupo de trabalhadores com ensino médio, todos os demais são menos demandados em função da mudança no padrão de gastos. Considerando todos os grupos, a demanda por trabalhadores com ensino superior completo é a que mais reduz em função da mudança de padrão de gastos.

Portanto, se a expansão dos gastos das famílias tende a demandar trabalho de baixa qualificação e, especialmente, daqueles ligados à fabricação de bens e serviços industriais e de reparação e manutenção, essa tendência se fortalece com a mudança de padrão ocorrida durante o período estudado.



Fonte: Resultados da pesquisa.

**Figura 5:** Impactos comparados: padrão de gastos das famílias de 2008 menos padrão de gastos das famílias de 2002 — por qualificação

#### 4 Considerações Finais

A partir da metodologia empregada, pode-se chegar a algumas conclusões importantes a respeito dos efeitos do crescimento dos gastos das famílias sobre a demanda por mão de obra. A primeira delas é relativa ao aumento de demanda pelo trabalho não qualificado. Especialmente quando a distribuição do impacto é comparada com a estrutura produtiva, percebe-se que os grupos de baixa qualificação — até ensino médio incompleto — recebem um percentual do impacto total maior que a participação do grupo na estrutura produtiva (embora o Grupo com ensino médio completo receba o maior impacto absoluto); enquanto os grupos de alta qualificação recebem um impacto proporcionalmente menor do que as respectivas participações no total da mão de obra empregada.

No caso das ocupações, a importância crescente da construção civil e a participação declinante dos alimentos no total das despesas parecem explicar os dois resultados mais expressivos: (1) o forte impacto na demanda pelos trabalhadores relativos aos bens e serviços industriais I e (2) o crescimento menor da demanda pelos trabalhadores agropecuários, florestais e da pesca do que teria sido caso não houvesse mudança no padrão dos gastos das famílias.

Portanto, em linhas gerais, a alteração ocorrida no padrão de consumo das famílias no Brasil tem impactado de forma mais significativa o emprego de baixa qualificação, especialmente ocupações ligadas à produção industrial.

#### Referências Bibliográficas

Barone, F. M. & Sader, E. (2008), 'Acesso ao crédito no Brasil: evolução e perspectivas', *Revista de Administração Pública* 42(6), 1249–1267.

Bielschowsky, R. (2012), 'Estratégia de desenvolvimento e as três frentes de expansão no Brasil: um desenho conceitual', *Economia e Sociedade* 21(número especial), 729–747.

Bonelli, R. & Cunha, P. V. (1981), 'Crescimento econômico, padrão do consumo e distribuição da renda no Brasil: uma abordagem multissetorial para

o período 1970/75', *Pesquisa e Planejamento Econômico* **11**(3), 703–756. Rio de Janeiro.

Bonelli, R. & Cunha, P. V. (1982), 'Mudanças nas estruturas de produção, renda e consumo, e crescimento econômico no Brasil no período 1970/75', *Pesquisa e Planejamento Econômico* **12**(3), 807–850. Rio de Janeiro.

Bonelli, R. & Cunha, P. V. (1983), 'Distribuição de renda e padrões de crescimento: um modelo dinâmico da economia brasileira', *Pesquisa e Planejamento Econômico* **13**(1), 91–154. Rio de Janeiro.

Bresser-Pereira, L. C. (2004), *Desenvolvimento e crise no Brasil*, Editora 34: São Paulo.

Cavalcanti, J. E. A. (1997), 'Distribuição setorial da renda: Seus efeitos de indução na economia brasileira', *Pesquisa e Planejamento Econômico* **27**(1), 141–184.

Cruyce, B. V. & Wera, J. (2007), 'Qualitative employment multipliers for Belgium, results for 2000 and 2002, in '16th International Input-Output Conference in Istanbul', Turquia.

Fochezatto, A. (2011), 'Estrutura da demanda final e distribuição de renda no Brasil: Uma abordagem multissetorial utilizando uma matriz de contabilidade social', *Economia (ANPEC)* **12**, 111–130. Brasília.

Guilhoto, J. J. M., Conceição, P. H. Z. & Crocomo, F. C. (1996), 'Estrutura de produção, consumo e distribuição de renda na economia brasileira: 1975 e 1980 comparados', *Economia & Empresa* **3**.

Jacinto, P. A. & Caetano, S. M. (2011), 'Os efeitos trabalhador adicional e desalento: uma análise para as regiões metropolitanas do nordeste', *Revista Econômica do Nordeste* **42**(2), 351–364.

Najberg, S. & Ikeda, M. (1999), 'Modelo de geração de emprego: metodologia e resultados, Texto para Discussão 72, BNDES, Rio de Janeiro.

Najberg, S. & Vieira, S. P. (1996), 'Modelos de geração de emprego aplicados à economia brasileira: 1985/95', *Revista do BNDES* **3**(5), 63–86. Rio de Janeiro.

Pyatt, G. (1988), 'A sam approach to modelling', *Journal of Policy Modeling* **10**(3), 327–352.

Pyatt, G. (2001), 'Some early multiplier models of the relationship between income distribution and production structure', *Economic Systems Research* **13**(2), 139–164.

Pyatt, G. & Round, J. I. (1977), 'Social accounting matrices for development planning', *Review of Income and Wealth Series* **23**(4), 339–364.

Pyatt, G. & Round, J. I. (1979), 'Accounting and fixed price multipliers in a sam framework', *The Economic Journal* **89**, 850–873.

Silva, T. L. F. & Moraes, G. I. (2013), 'Determinantes da demanda por crédito pessoa física — recursos livres, no período 2004-2011, em estados selecionados', *Pesquisa & Debate* **23**(2), 327–344.

Stone, J. R. N. (1985), The disaggregation of the household sector in the national accounts, *in* G. Pyatt & J. I. E. Round, eds, 'Social Accounting Matrices: A Basis for Planning', The World Bank, Washington D.C, pp. 145–185.

Thorbecke, E. (1995), Intersectoral linkages and their impact on rural poverty alleviation: A social accounting approach, Technical report, United Nations Development Organisation (UNIDO), Vienna.

Torres, E. T. (2010), Crédito cresce forte em 2009, apesar da crise internacional. visão do desenvolvimento, Technical Report 80, BNDES.

## **Apêndice A**

**Tabela A.1:** Número de postos de trabalho gerados por setor e por qualificação

Setores	A	B	C	D	E	F	G	H	I	J	K	Total
Agricultura, silvicultura, exploração flo- restal	104955	536542	540305	515670	401616	168060	363394	25578	61659	1370	587	2719734
Pecuária e pesca	38385	182747	194049	190229	171120	70545	137121	8550	23187	350	86	1016370
Petróleo e gás natural	1	46	18	55	219	108	3659	357	2868	9	2	7341
Minério de ferro	2	20	19	30	83	62	885	61	303	4	0	1469
Outros da indústria extrativa	342	2496	3483	4873	5753	2651	10027	717	1841	26	20	32230
Alimentos e bebidas	12046	48084	41970	70923	85621	56897	180929	14580	21921	331	60	533361
Produtos do fumo	13	110	168	206	329	309	1826	401	647	6	2	4018
Têxteis	542	6747	16548	32670	50574	35948	109646	7154	12014	151	21	272015
Artigos do vestuário e acessórios	2946	18800	51821	127308	241255	156727	429651	20022	19936	216	35	1068718
Artefatos de couro e calçados	168	1819	3672	13127	10647	9299	18698	1163	794	8	1	59395
Produtos de madeira — exclusive mó- veis	871	5439	8207	12862	14382	6749	17054	1027	1512	13	4	68118
Celulose e produtos de papel	53	623	1273	2375	4242	2666	13338	1237	2521	30	8	28366
Jornais, revistas, discos	41	549	1172	3117	7072	5960	29675	4411	9134	79	12	61224
Refino de petróleo e coque	7	78	144	253	425	259	1748	191	619	4	1	3728
Álcool	1048	6473	4743	5874	4388	2701	7209	620	1420	12	3	34491
Produtos químicos	78	585	868	1867	2344	1305	8710	1797	5101	105	19	22778
Fabricação de resina e elastômeros	1	59	120	262	686	369	2956	490	1197	22	6	6168
Produtos farmacêuticos	54	693	1184	2414	4021	4986	38521	10242	27433	248	168	89964
Defensivos agrícolas	5	74	103	208	240	200	1319	212	1132	18	4	3515
Perfumaria, higiene e limpeza	105	1393	1785	3823	6507	5471	28161	2426	4744	51	14	54480
Tintas, vernizes, esmaltes e lacas	9	128	198	382	836	583	3113	535	963	17	11	6776
Produtos e preparados químicos diver- sos	60	670	880	1237	2106	1224	7494	1243	3786	74	34	18808
Artigos de borracha e plástico	216	2056	4239	10368	18382	11722	53437	3810	5763	47	10	110049
Cimento	9	80	92	155	216	192	1388	164	458	3	1	2758
Outros prod. de minerais não metálicos	1557	8460	12977	18911	24337	11710	33561	2400	3962	40	6	117921
Fabricação de aço e derivados	33	536	822	1440	2083	1522	11938	886	3180	64	16	22521
Metalurgia de metais não ferrosos	23	315	691	1332	2100	992	5473	453	879	7	1	12265

A = Analfabeto, B = Até 5ª Incompleto, C = 5ª Completo Fundamental, D = 6ª a 9ª Fundamental, E = Fundamental Completo, F = Médio Incompleto, G = Médio Completo, H = Superior Incompleto, I = Superior Completo, J = Mestrado, K = Doutorado.

Fonte: Resultados da pesquisa.

**Tabela A.1:** Número de postos de trabalho gerados por setor e por qualificação (continuação)

Setores	A	B	C	D	E	F	G	H	I	J	K	Total
Prod de metal — exclusive máq e equipamentos	517	3963	8159	18302	35392	20943	79431	6058	8106	96	16	180983
Máq. e equip., inclusive manutenção e reparos	42	469	1097	2378	4840	3193	16064	1968	3292	41	12	33397
Eletrodomésticos	176	622	1043	2670	5741	3782	35006	1982	5648	67	17	56754
Máq. para escritório e equip. de informática	0	2	3	12	29	47	616	79	136	2	0	925
Máquinas, aparelhos e materiais elétricos	59	385	716	1984	3899	3436	17636	1857	3888	40	5	33906
Material eletrônico e equip. de comunicações	12	81	172	430	961	978	8842	634	1455	25	2	13592
Apar./instrum. médico-hosp., medida e óptico	37	288	618	1558	2751	1941	8406	1001	1649	32	17	18298
Automóveis, camionetas e utilitários	8	481	1200	1144	7458	4142	23721	3604	14049	81	16	55904
Caminhões e ônibus	-	2	4	12	22	24	180	31	103	0	0	378
Peças e acessórios para veículos automotores	224	1779	3921	7825	17969	11554	75073	5783	11107	156	20	135411
Outros equipamentos de transporte	5	88	162	357	610	389	2648	179	612	5	1	5057
Móveis e produtos das indústrias diversas	998	6769	14533	32911	48739	29517	83175	6189	6737	68	29	229665
Elétric. e gás, água, esgoto e limpeza urbana	282	3419	3984	8076	8383	4903	34756	4504	18357	872	69	87606
Construção	15233	141380	163196	261615	298511	117858	361462	24175	58667	771	299	1443167
Comércio	2273	18045	36969	88421	193136	162730	752838	55886	68761	551	323	1379932
Transporte, armazenagem e correio	1023	16426	43151	84277	139300	70746	325682	25441	52263	341	62	758712
Serviços de informação	97	938	1629	4053	10710	10586	110965	45658	103770	742	375	289521
Intermediação financeira e seguros	137	919	1689	3461	6732	4298	32141	13975	38913	270	44	102578
Serviços imobiliários e aluguel	972	8075	10743	18515	29055	15163	70701	8734	13340	127	26	175452
Serviços de manutenção e reparação	7112	26250	57144	135949	294443	191472	723404	58971	63700	1365	295	1560105

A = Analfabeto, B = Até 5ª Incompleto, C = 5ª Completo Fundamental, D = 6ª a 9ª Fundamental, E = Fundamental Completo, F = Médio Incompleto, G = Médio Completo, H = Superior Incompleto, I = Superior Completo, J = Mestrado, K = Doutorado.

Fonte: Resultados da pesquisa.

**Tabela A.1:** Número de postos de trabalho gerados por setor e por qualificação (continuação)

Setores	A	B	C	D	E	F	G	H	I	J	K	Total
Serviços de alojamento e alimentação	3644	17 662	43 493	87 506	184 370	111 873	317 950	12 497	20 406	226	35	799 663
Serviços prestados às empresas	1699	11 695	21 540	42 729	78 472	39 362	272 347	33 947	52 759	862	607	556 020
Educação mercantil	-3	-23	-33	-65	-105	-69	-582	-145	-1209	-95	-48	-2375
Saúde mercantil	386	3354	6337	11 683	23 705	18 403	175 104	13 088	62 274	386	201	314 921
Outros serviços	11 199	77 647	91 689	132 333	198 852	119 487	389 753	46 355	160 708	3790	971	1 232 786
Educação pública	5	40	47	58	140	48	454	55	463	6	1	1316
Saúde pública	0	1	1	1	3	1	9	1	10	0	0	27
Administração pública e seguridade social	81	715	853	1040	2533	873	8209	986	8365	115	14	23786
<b>Total</b>	<b>209788</b>	<b>1 167 094</b>	<b>1 405 611</b>	<b>1 971 206</b>	<b>2 658 235</b>	<b>1 506 897</b>	<b>5 446 922</b>	<b>484 220</b>	<b>997 303</b>	<b>14 247</b>	<b>4541</b>	<b>15 866 069</b>
Percentual do total	1,32%	7,36%	8,86%	12,42%	16,75%	9,50%	34,33%	3,05%	6,29%	0,09%	0,03%	100,00%

A = Analfabeto, B = Até 5ª Incompleto, C = 5ª Completo Fundamental, D = 6ª a 9ª Fundamental, E = Fundamental Completo, F = Médio Incompleto, G = Médio Completo, H = Superior Incompleto, I = Superior Completo, J = Mestrado, K = Doutorado.

Fonte: Resultados da pesquisa.

**Tabela A.2:** Número de postos de trabalho gerados por setor e por ocupação

Setores	Ocup1	Ocup2	Ocup3	Ocup4	Ocup5	Ocup6	Ocup7	Ocup8	Ocup9	NC	Total
Agricultura, silvicultura, exploração florestal	38092	31389	60912	114347	121109	2026399	249725	23732	53736	293	2719734
Pecuária e pesca	20482	14102	18461	41957	59649	775080	59487	18650	8457	46	1016370
Petróleo e gás natural	134	2068	1712	755	155	6	1308	1016	186	0	7341
Minério de ferro	53	145	316	81	21	1	528	19	304	-	1469
Outros da indústria extrativa	751	995	2015	2889	2369	389	18707	1553	2559	4	32230
Alimentos e bebidas	13347	8810	25421	51598	89037	70071	106791	153649	14609	28	533361
Produtos do fumo	274	251	826	410	93	94	785	902	383	-	4018
Têxteis	5695	5068	20993	26243	17605	836	170891	12321	12350	15	272015
Artigos do vestuário e acessórios	23980	11705	35848	83510	95254	639	801340	8280	8131	30	1068718
Artefatos de couro e calçados	743	421	1656	3214	1240	55	50083	1467	516	1	59395
Produtos de madeira — exclusive móveis	1751	708	1863	5253	3468	3505	47532	2411	1624	3	68118
Celulose e produtos de papel	861	908	2620	2907	1353	357	10594	7550	1208	7	28366
Jornais, revistas, discos	2400	7173	5408	13283	5617	43	25517	1173	604	6	61224
Refino de petróleo e coque	85	301	555	401	294	35	1075	836	146	1	3728
Álcool	322	679	1590	2137	1516	16177	6598	3295	2177	-	34491
Produtos químicos	1515	1932	3154	3534	1235	452	4894	4905	1157	-	22778
Fabricação de resina e elastômeros	265	491	1295	623	167	6	1331	1748	241	1	6168
Produtos farmacêuticos	5929	7605	22718	12246	15705	121	10382	12805	2347	106	89964
Defensivos agrícolas	307	351	627	364	215	262	643	639	107	-	3515
Perfumaria, higiene e limpeza	1956	1796	4575	8358	8063	131	19553	8180	1850	16	54480
Tintas, vernizes, esmaltes e lacas	328	367	1139	1141	522	11	1322	1806	140	-	6776
Produtos e preparados químicos diversos	1123	1356	2599	2667	1373	461	3982	4568	677	1	18808
Artigos de borracha e plástico	2850	2313	9933	11890	5761	211	42987	28901	5196	6	110049
Cimento	105	164	510	351	153	5	790	249	431	-	2758

Ocup. 1 = Membros superiores do poder público, dirigentes de organizações de interesse público, Ocup. 2 = Profissionais das ciências e das artes, Ocup. 3 = Técnicos de nível médio, Ocup. 4 = Trabalhadores de serviços administrativos, Ocup. 5 = Trabalhadores dos serviços, vendedores do comércio em lojas e mercados, Ocup. 6 = Trabalhadores agropecuários, florestais e da pesca, Ocup. 7 = Trabalhadores da produção de bens e serviços industriais I, Ocup. 8 = Trabalhadores da produção de bens e serviços industriais II, Ocup. 9 = Trabalhadores em serviços de reparação e manutenção, NC = Não Classificado.

Fonte: Resultados da pesquisa.

**Tabela A.2:** Número de postos de trabalho gerados por setor e por ocupação (continuação)

Setores	Ocup1	Ocup2	Ocup3	Ocup4	Ocup5	Ocup6	Ocup7	Ocup8	Ocup9	NC	Total
Outros produtos de minerais não metálicos	2695	1766	4567	9498	6906	481	61951	26592	3455	10	117921
Fabricação de aço e derivados	537	1613	3658	1950	693	186	7268	4476	2139	0	22521
Metalurgia de metais não ferrosos	283	391	1255	870	356	16	6585	1714	794	0	12265
Produtos de metal — exclusive máquinas e equipamentos	4020	3651	13765	17963	8620	293	113951	11018	7695	6	180983
Máquinas e equipamentos, inclusive manutenção e reparos	1005	1423	4498	3825	1585	58	16615	1529	2858	1	33397
Eletrodomésticos	852	2445	5070	4653	2506	32	35745	4269	1181	-	56754
Máquinas para escritório e equipamentos de informática	33	74	147	156	46	0	437	18	12	0	925
Máquinas, aparelhos e materiais elétricos	1042	1957	4015	3546	1084	20	18803	2061	1375	1	33906
Material eletrônico e equipamentos de comunicações	440	821	2281	1661	581	10	7234	296	266	1	13592
Aparelhos/instrumentos médico-hospitalar, medida e óptico	596	819	2514	2442	980	18	9576	728	624	0	18298
Automóveis, camionetas e utilitários	1520	5989	6053	2865	909	-	33922	329	4317	-	55904
Caminhões e ônibus	7	35	69	24	7	0	207	2	28	-	378
Peças e acessórios para veículos automotores	2653	5006	14223	11357	2870	72	83494	9450	6282	4	135411
Outros equipamentos de transporte	114	357	641	465	127	3	2677	212	463	0	5057
Móveis e produtos das indústrias diversas	5900	2584	8767	23275	15477	544	158056	11602	3443	19	229665

Ocup. 1 = Membros superiores do poder público, dirigentes de organizações de interesse público, Ocup. 2 = Profissionais das ciências e das artes, Ocup. 3 = Técnicos de nível médio, Ocup. 4 = Trabalhadores de serviços administrativos, Ocup. 5 = Trabalhadores dos serviços, vendedores do comércio em lojas e mercados, Ocup. 6 = Trabalhadores agropecuários, florestais e da pesca, Ocup. 7 = Trabalhadores da produção de bens e serviços industriais I, Ocup. 8 = Trabalhadores da produção de bens e serviços industriais II, Ocup. 9 = Trabalhadores em serviços de reparação e manutenção, NC = Não Classificado.

Fonte: Resultados da pesquisa.

**Tabela A.2:** Número de postos de trabalho gerados por setor e por ocupação (continuação)

Setores	Ocup1	Ocup2	Ocup3	Ocup4	Ocup5	Ocup6	Ocup7	Ocup8	Ocup9	NC	Total
Eletricidade e gás, água, esgoto e limpeza urbana	4320	9886	12049	16181	13049	190	19238	10004	2684	5	87606
Construção	16004	37151	69220	111903	114685	7477	1007070	11006	68552	101	1443167
Comércio	78285	33942	67745	332850	567768	7816	199970	44263	47143	150	1379932
Transporte, armazenagem e correio	15361	17335	50135	159312	144365	5808	333566	3861	28912	57	758712
Serviços de informação	15608	86778	69072	82356	18040	145	13731	808	2960	23	289521
Intermediação financeira e seguros	8281	12170	8589	50906	15331	349	5288	476	1182	6	102578
Serviços imobiliários e aluguel	6543	6423	10902	44408	40224	1753	58484	980	5706	29	175452
Serviços de manutenção e reparação	48773	44061	174134	268160	264568	2891	339117	12515	405646	239	1560105
Serviços de alojamento e alimentação	40820	7694	11013	100291	582616	2539	19159	30268	5146	117	799663
Serviços prestados às empresas	9775	32617	40845	170782	233460	4069	49670	5331	9358	113	556020
Educação mercantil	-55	-1168	-441	-397	-255	-6	-39	-4	-10	0	-2375
Saúde mercantil	5023	51496	102722	86393	59127	510	6997	852	1771	29	314921
Outros serviços	28538	130863	113784	246034	566780	11762	116913	5171	12766	174	1232786
Educação pública	118	308	262	260	214	4	44	3	13	91	1316
Saúde pública	2	6	5	5	4	0	1	0	0	2	27
Administração pública e seguridade social	2131	5564	4733	4700	3877	65	799	48	227	1643	23786
<b>Total</b>	<b>424577</b>	<b>605155</b>	<b>1033042</b>	<b>2148856</b>	<b>3098573</b>	<b>2942452</b>	<b>4363376</b>	<b>500531</b>	<b>746121</b>	<b>3386</b>	<b>15866069</b>
Percentual do total	2,68%	3,81%	6,51%	13,54%	19,53%	18,55%	27,50%	3,15%	4,70%	0,02%	100,00%

Ocup. 1 = Membros superiores do poder público, dirigentes de organizações de interesse público, Ocup. 2 = Profissionais das ciências e das artes, Ocup. 3 = Técnicos de nível médio, Ocup. 4 = Trabalhadores de serviços administrativos, Ocup. 5 = Trabalhadores dos serviços, vendedores do comércio em lojas e mercados, Ocup. 6 = Trabalhadores agropecuários, florestais e da pesca, Ocup. 7 = Trabalhadores da produção de bens e serviços industriais I, Ocup. 8 = Trabalhadores da produção de bens e serviços industriais II, Ocup. 9 = Trabalhadores em serviços de reparação e manutenção, NC = Não Classificado.

Fonte: Resultados da pesquisa.



## EFEITOS DOS PISOS SALARIAIS ESTADUAIS SOBRE O MERCADO DE TRABALHO: UMA NOVA ABORDAGEM EMPÍRICA

CARLOS HENRIQUE CORSEUIL \*

MIGUEL FOGUEL †

MARCOS HECKSHER ‡

### Resumo

Este trabalho avalia os efeitos dos pisos salariais introduzidos no Paraná e em São Paulo sobre alguns indicadores do mercado de trabalho. Em ambas as Unidades da Federação (UFs), são analisados três grupos ocupacionais distintos com base nos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD/IBGE). Conclui-se que, em um dos grupos de cada UF, a política aumentou a remuneração de trabalhadores que receberiam menos que os respectivos pisos caso estes não tivessem sido criados, sem se contrair o emprego ou se aumentar a informalidade. Nos demais agregados ocupacionais de cada UF, não foi encontrado qualquer efeito significativo da legislação. A maior contribuição deste trabalho é a utilização de uma nova metodologia mais apropriada para identificar efeitos de tratamento em unidades agregadas, apresentada por Abadie et al. (2010).

**Palavras-chave:** pisos salariais; salário mínimo; mercado de trabalho; efeito de tratamento.

### Abstract

This paper evaluates the introduction of Brazilian regional and occupational wage floors on labor market performance. We focus on two out of five states that implemented such intervention: Paraná and São Paulo. In each state, we have analyzed the effects of the wage floors for three distinct occupational categories. Our results point that one occupational group in each state seems to be affected by the introduction of a wage floor. In both cases, the wage floor tends to reduce the share of employees with wages below the established floor without producing side effect on employment or on informality. It was not found any significant effect of the new legislation in the other occupational groups in each state. We innovate on methodological grounds by using the synthetic control method put forward by Abadie et al. (2010).

**Keywords:** wage floors; minimum wages; labor market; treatment effect.

**JEL classification:** J21; J31; J38

**DOI:** <http://dx.doi.org/10.1590/1413-8050/ea125210>

\* IPEA. E-mail: carlos.corseuil@ipea.gov.br

† IPEA. E-mail: miguel.foguel@ipea.gov.br

‡ IPEA e mestrando em População, Território e Estatísticas Públicas da Ence/IBGE. E-mail: marcos.hecksher@ipea.gov.br

## 1 Introdução

Pisos salariais legislados em valores acima do salário mínimo (SM) nacional estão em vigor atualmente em cinco estados brasileiros: Rio de Janeiro (RJ), Rio Grande do Sul (RS), Paraná (PR), São Paulo (SP) e Santa Catarina (SC). Em dois destes estados (RJ e RS), os pisos vigoram há mais de dez anos. No entanto, ainda há poucas avaliações sobre os efeitos dessas medidas em indicadores do mercado de trabalho. Um dos grandes obstáculos para que essa tarefa seja executada é encontrar uma estratégia de identificação que seja ao mesmo tempo convincente e aplicável a um conjunto amplo de indicadores. Uma das principais contribuições deste estudo vem justamente nessas dimensões, ao utilizar uma estratégia ainda inédita nesse tipo de análise.

A estratégia empírica consagrada na literatura internacional para identificar o efeito de pisos salariais estaduais sobre indicadores do mercado de trabalho é o método de diferenças em diferenças.<sup>1</sup> Para que essa estratégia seja bem-sucedida, supõe-se que a evolução dos indicadores do mercado de trabalho nos estados que criaram ou aumentaram seus pisos salariais seja bem aproximada pela evolução desses indicadores em um conjunto de estados no período pós-intervenção.<sup>2</sup> Em outras palavras, usa-se a evolução nesse conjunto de estados após a intervenção como uma aproximação do que seria visto nas unidades que sofreram a intervenção caso esta não tivesse sido realizada. Naturalmente, esta suposição não é diretamente testável, mas em algumas aplicações utiliza-se um indicativo indireto da sua validade baseado em testar se o indicador de interesse tem a mesma evolução pré-intervenção.<sup>3</sup> Em aplicações desse método, também é comum o uso apenas de um estado vizinho para desempenhar o papel de contrafactual.

Este texto emprega, pela primeira vez nesse tipo de análise, a estratégia de identificação proposta por Abadie & Gardeazabal (2003) e posteriormente formalizada por Abadie et al. (2010). Nessa estratégia, a construção de um contrafactual da evolução do indicador de interesse para o grupo tratado procura combinar informações de todos os demais grupos não tratados. Essa combinação é obtida por intermédio de uma média ponderada em que os pesos visam minimizar a distância entre os dois grupos no período pré-intervenção. Isso torna essa estratégia mais flexível que a de diferenças em diferenças, uma vez que, por construção, ela procura “sintetizar” um grupo de controle similar ao grupo tratado antes da intervenção.<sup>4</sup> Ademais, como mostra a seção 4, o modelo por trás da estratégia proposta por Abadie et al. (2010) é mais geral que o de diferenças em diferenças.

---

<sup>1</sup>Card (1992) e, principalmente, Card & Krueger (1994) podem ser apontados como os principais responsáveis pela consagração dessa estratégia no contexto de avaliações de pisos salariais por dois motivos. Em primeiro lugar, foram pioneiros no uso desta metodologia na literatura de SM. Em segundo, por serem trabalhos que, juntamente com Card & Krueger (1995), tiveram grande repercussão no debate em torno dos efeitos do SM no mercado de trabalho.

<sup>2</sup>Mais detalhes dessa estratégia podem ser obtidos em Meyer (1995) ou em resenhas metodológicas tais como Angrist & Krueger (1999) ou BLUNDELL (2009).

<sup>3</sup>No entanto, em muitos casos, esse teste não é feito e a análise é conduzida com a possibilidade de que os grupos de tratamento e de controle sequer tenham tido a mesma trajetória antes da intervenção.

<sup>4</sup>Uma estratégia alternativa, adotada na literatura visando aumentar as chances de o contrafactual reproduzir a evolução do grupo afetado pela política, consiste em empregar o método de diferenças em diferenças restringindo a amostra a áreas perto da divisa entre o estado que sofreu a intervenção e o que não sofreu a intervenção. Este tipo de estratégia é utilizado em Dube et al. (2010) Dube et al. (2011).

Neste estudo, essa estratégia de identificação é empregada para avaliar o efeito dos pisos salariais sobre o mercado de trabalho de dois dos cinco estados que adotaram a política, a saber, PR e SP. Essa escolha se deve ao fato de que os pisos são fixados por grupos ocupacionais definidos pelos próprios estados e somente para esses dois foi possível compatibilizar as listas de ocupações com as da PNAD, a base de dados aqui utilizada. Além disso, houve uma mudança na classificação das ocupações da PNAD a partir de 2002 e para evitar a necessidade de uma segunda compatibilização das ocupações — que poderia reduzir a acurácia dos resultados —, optou-se por utilizar somente as versões da PNAD entre 2002 e 2009. Como a metodologia aqui utilizada requer informações antes da intervenção, somente o PR e SP atendem a condição de terem instituído seus pisos durante esse intervalo.

O método será usado para avaliar o impacto dos pisos salariais nas seguintes variáveis: proporção de empregados com remuneração inferior ao valor do piso de referência; taxa de emprego; taxa de informalidade; e hiato salarial entre empregados com e sem carteira de trabalho assinada. Julga-se ser do interesse do legislador que o piso eleve a remuneração dos trabalhadores que ocupam postos remunerados abaixo do que foi legislado, o que justifica a escolha da primeira variável a ser analisada. No entanto, para que isso efetivamente ocorra, os agentes que interagem no mercado de trabalho devem reagir a um aumento de piso salarial tal como desejado pelo poder público, o que nem sempre ocorre.

A escolha dos demais indicadores foi guiada pelo debate na literatura a respeito da maneira como o mercado de trabalho se ajusta a um aumento de pisos salariais, ou seja, por meio de ajustes no nível de emprego, na sua composição e nos salários praticados nos segmentos formal e informal do mercado de trabalho.

De acordo com o arcabouço teórico clássico sobre SM, as firmas devem reagir ao aumento do piso diminuindo o nível de emprego. Card & Krueger (1994) Card & Krueger (1995) contestaram esse resultado mostrando evidências que indicam um efeito nulo ou um pequeno aumento do emprego em decorrência de aumentos de pisos salariais estaduais nos Estados Unidos.<sup>5</sup> Posteriormente, associaram-se esses efeitos nulos ou positivos no emprego a predições teóricas dos modelos monopsonísticos tradicionais ou de busca por emprego (*search models*).<sup>6,7</sup> Esse debate justifica o interesse na taxa de emprego.

O citado debate ignora dois pontos relevantes sobre o ajuste do mercado de trabalho à lei dos pisos. O primeiro é a presença do setor informal e o segundo é a possibilidade de cumprimento parcial da lei. Com relação ao primeiro ponto, há indícios de que o ajuste pode ser distinto quando os trabalhadores têm duas alternativas para se empregarem: em um posto formal ou num posto informal. Do ponto de vista teórico, Mincer (1976) trata desse

<sup>5</sup>Essas evidências foram contestadas por Neumark & Wascher (2000), que argumentam que elas são dependentes dos dados utilizados por Card & Krueger (1994). Roppnen (2011) mostra que, na verdade, o impacto depende do tamanho da firma e, portanto, a diferença de resultados reportados neste debate se devia à diferença na composição das respectivas amostras no que diz respeito ao tamanho das firmas. Debates similares ocorreram em outros países — ver, por exemplo, Garloff (2010) para o caso alemão.

<sup>6</sup>O leitor interessado nessa associação pode consultar Swinnerton (1996) e Masters (1999). Garloff (2010) mostra que, mesmo dentro da classe de modelos de busca, há uma ambiguidade sobre a direção do efeito de pisos salariais sobre o emprego.

<sup>7</sup>Há outras vertentes teóricas capazes de racionalizar o resultado de Card & Krueger (1994) Card & Krueger (1995) o que é comentado em mais detalhes na subseção 5.4.

tema adaptando o arcabouço clássico para incorporar esses dois setores, e mostra que pode haver migração de trabalhadores do setor formal para o informal após a introdução ou a elevação do SM.<sup>8</sup> Empiricamente, há diversos trabalhos que apresentam evidências de que o setor informal tende a ser afetado pelo SM. Na realidade, essa influência pode se dar também de forma direta, uma vez que as firmas podem decidir cumprir a legislação sobre os pisos em seus postos informais.<sup>9</sup> Esses efeitos, diretos ou indiretos, sobre níveis de emprego e salário dos setores formal e informal são um dos motivadores da escolha do indicador de hiato salarial entre empregados com e sem carteira.

Com relação ao segundo ponto, Danziger (2010) reporta que há evidências de cumprimento parcial em pisos salariais para vários países, inclusive para os Estados Unidos. Moura & Neri (2008) reportam evidências na mesma direção nos primeiros anos de vigência dos pisos estaduais no RJ e no RS. A maneira mais intuitiva de considerar esse problema pode ser exemplificada a seguir. Eventualmente, o piso estadual pode ter efeitos sobre os trabalhadores quando estes estão, de fato, sujeitos à lei. No entanto, se uma grande parcela dos empregadores não cumpre a lei, é de se esperar que tal efeito seja pulverizado em uma análise que agregue tanto os trabalhadores de firmas cumpridoras quanto os de firmas descumpridoras, que mantêm empregados com remuneração menor que o piso. Além disso, o não cumprimento pode alterar de forma menos intuitiva a maneira com que o mercado de trabalho se ajusta (Yaniv 2004, Danziger 2009).

Não há muitas alternativas para contornar o problema de cumprimento parcial dos pisos salariais nos estados brasileiros. O que se faz aqui é apenas inferir em que medida o grau de cumprimento interfere na análise. Como os pisos salariais foram definidos por grupos de ocupação em cada Unidade da Federação (UF) que os adotou, foram conduzidas análises agregadas por UF e também análises estratificadas por grupos ocupacionais com diferentes taxas de cumprimento em cada UF. Dessa forma, é possível verificar se eventuais efeitos para grupos com cumprimento relativamente maior são pulverizados no agregado da UF.

Cabe destacar que uma avaliação dos efeitos dos pisos estaduais sobre indicadores do mercado de trabalho também pode ser útil para uma avaliação dos efeitos do SM nacional, em que a tarefa de encontrar uma estratégia de identificação adequada seria ainda mais árdua. Afinal, o piso nacional é homogêneo em todos os estados, todas as ocupações e todos os setores. A experiência dos pisos estaduais permite ao analista fazer uso de heterogeneidades que não podem ser vistas no mínimo nacional. Além disso, essa experiência por estado ainda permite isolar efeitos diretos sobre o mercado de trabalho daqueles intermediados por transferências públicas de renda, indexadas somente ao SM nacional.

Apesar da utilidade das avaliações dos efeitos dos pisos para mensurar os impactos do SM, há pelo menos três cuidados a serem tomados. Em primeiro

---

<sup>8</sup>Além desses dois setores, o modelo de Mincer (1976) permite a existência do estado de desemprego. Assim, dependendo do hiato salarial que aparece entre os setores formal e informal, é possível que os trabalhadores prefiram permanecer desempregados à espera de uma vaga no setor formal.

<sup>9</sup>Entre os diversos trabalhos mais recentes que se debruçaram sobre a relação entre o SM e o setor informal no Brasil, ver Fajnzylber (2001), Neri et al. (2001), Soares (2004) e Lemos (2004). Resenhas sobre os impactos do SM no mercado de trabalho no Brasil podem ser encontradas em Corseuil & Servo (2002) e Ulyssea & Foguel (2006). Para uma resenha dos efeitos do SM na América Latina, ver Cunningham (2007).

lugar, se os efeitos do SM sobre indicadores de mercado de trabalho variam dependendo do valor do mínimo, os impactos estimados com base nos pisos estaduais podem fornecer uma resposta equivocada para os efeitos do mínimo. Em segundo lugar, se os impactos do mínimo mudam de acordo com o seu grau de cumprimento no mercado de trabalho, como os pisos têm diferentes graus de cumprimento locais relativamente ao do SM, os efeitos computados com o uso dos pisos podem ser pouco informativos acerca dos efeitos do mínimo. Finalmente, como a natureza e a magnitude dos impactos captados com base nos pisos salariais de uma UF podem ser específicas daquela UF, a generalização desses impactos para o âmbito nacional pode ser arriscada.

Este trabalho está organizado em mais cinco seções. A seção 2 descreve o contexto institucional da intervenção analisada, com detalhes relevantes da lei federal que introduziu a possibilidade de diferenciação de pisos salariais nas UFs, bem como aspectos específicos da legislação dos pisos em algumas UFs. Em seguida, a seção 3 traz uma descrição dos dados acompanhada por algumas estatísticas descritivas. A seção 4 apresenta a metodologia e a seção 5, os resultados. Por fim, a seção 6 apresenta as conclusões.

## 2 A legislação dos pisos estaduais no Brasil e suas avaliações

A variedade de pisos salariais legais em diferentes estados brasileiros foi criada a partir de uma autorização expressamente incluída na legislação federal em 2000. Embora a Constituição Federal (Brasil 1988) mantenha o SM nacionalmente unificado,<sup>10</sup> também prevê que se estabeleça um piso salarial “proporcional à extensão e à complexidade do trabalho” (Brasil, 1988, Artigo 7º, inciso V). Com base nisso, desde 2000, a Lei Complementar (LC) nº103 permite que os governos dos estados e do Distrito Federal proponham leis com pisos salariais superiores ao mínimo nacional. Os pisos estaduais devem regular ocupações ou atividades específicas e só valem na ausência de leis federais ou negociações coletivas que fixem outros pisos — superiores ou inferiores aos estaduais, mas nunca menores que o mínimo nacional. É proibido exigir que as administrações municipais cumpram os pisos estaduais. Nem os próprios governos estaduais são obrigados a segui-los, o que limita seu impacto fiscal.<sup>11</sup>

Após a lei federal de 2000, cinco estados passaram a fixar pisos acima do SM: RJ e RS, desde 2001; PR, desde 2006; SP, desde 2007; e SC, desde 2010. Os cinco pertencem às regiões Sul e Sudeste do país, onde já se pagavam salários mais altos que a média nacional, mas há uma grande diversidade de casos. Os valores, as datas de reajuste, as quantidades de faixas e as categorias profissionais ou atividades empresariais que compõem cada faixa variam bastante no espaço e no tempo. O RJ, por exemplo, começou com três faixas diferentes e foi aumentando até chegar às nove atuais, com muitas mudanças nas ocupações inseridas em cada uma. O PR iniciou com cinco faixas e reduziu para três.

<sup>10</sup>O SM, fixado pelo governo federal, foi criado no Brasil em 1940 com quatorze valores diferenciados entre estados e sub-regiões. Em 1963, chegou a ter 38 faixas regionais (Saboia 1984). Sua unificação nacional foi estabelecida em 1984.

<sup>11</sup>O RS, desde 2003, decidiu pagar também um complemento a seus servidores ativos e inativos com remuneração inferior ao piso mais alto do estado. Mas, como a lei federal não obriga nenhum estado a fazer isso, a regra geral é que os pisos estaduais só precisam ser pagos pelo setor privado, enquanto os governos estaduais que os propõem não incorrem em custos diretos.

Além de dividir com os governadores parte da pressão por aumentos no SM, um dos propósitos aparentes da lei federal de 2000 era desvincular os pisos do setor privado daqueles que pesam sobre salários e benefícios pagos pelo setor público nas três esferas. No debate sobre políticas para o mercado de trabalho anterior à lei, este era um dos argumentos de Jatobá & Chahad (1997) e de Neri (1999) em defesa da medida, que faria parte de uma agenda politicamente viável em alternativa à desvinculação explícita entre benefício previdenciário e SM. Em geral, a desvinculação costuma ser proposta para permitir que o SM possa aumentar sem pressionar as contas públicas, enquanto o piso previdenciário passa a ser indexado a algum indicador que tenda a subir mais lentamente. Na proposta de regionalização, resultado semelhante era buscado por um caminho quase inverso: o SM passaria a definir apenas remunerações e benefícios pagos pelo setor público e seria reajustado no ritmo que a política fiscal permitisse, ao passo que os pisos estaduais poderiam subir mais e passar a ser o foco de pressão dos trabalhadores ativos do setor privado, que os negociariam com entidades patronais e governos de seus estados.

Passados quinze anos, observa-se que a valorização real do SM foi intensificada<sup>12</sup> e a regra de reajuste vigente desde 2008, que acarretou altas superiores às do produto médio por trabalhador, está perto de ter sua extensão até 2019 fixada em nova lei.<sup>13</sup> Enquanto isso, nos estados que instituíram pisos, os salários pagos pelo setor privado sempre aderiram bem menos aos pisos estaduais que ao mínimo nacional, embora a fiscalização de ambos esteja a cargo do Ministério do Trabalho e Emprego (MTE). Dois outros estudos, descritos a seguir, constataram este alto descumprimento de pisos estaduais.

Por meio de regressões quantílicas com microdados longitudinais de 2000 a 2001 da Pesquisa Mensal de Emprego (PME/IBGE), Moura & Neri (2008) observaram um baixo cumprimento dos pisos logo após sua criação no RJ e no RS. Tomando os trabalhadores de SP como grupo de controle, os autores encontraram efeito estatisticamente nulo da política sobre os níveis de salário e de emprego nos dois estados que a adotaram em 2001.

Com base em microdados de 2006 a 2008 da Pesquisa Nacional por Amostra de Domicílios (PNAD/IBGE), Chahad & Pereda (2010) investigaram a efetividade das três faixas de pisos vigentes em SP em 2007 e 2008. Utilizando Minas Gerais (MG) como grupo de controle em um modelo *logit* ordenado, os autores concluíram que a lei podia ser considerada efetiva nas duas primeiras faixas salariais, embora o cumprimento não fosse alto.

### 3 Dados

Para avaliarem-se os graus de cumprimento e descumprimento dos pisos e seus efeitos sobre indicadores do mercado de trabalho, o primeiro passo é saber a quem se aplica cada valor fixado nas leis estaduais. Essa questão, no entanto, só admite uma resposta inequívoca no PR e a partir de 2007, pois, desde então, os pisos paranaenses passaram a ser explicitamente associados

---

<sup>12</sup>A média móvel em 12 meses anteriores do salário mínimo real subiu 4,0% ao ano entre julho de 1995 e julho de 2000, quando foi publicada a lei federal que autorizou os pisos estaduais. Desde então, a alta média foi de 4,9% ao ano.

<sup>13</sup>Pela regra adotada desde 2008 e fixada na Lei no 12.382/2011, o SM recebe, a cada ano, um aumento real igual ao último dado anual disponível de crescimento real do produto interno bruto (PIB). Até 2014, como a população ocupada cresceu inclusive nos anos de crise, o PIB cresceu mais que o PIB por trabalhador.

na lei a grandes grupos (GG) da classificação brasileira de ocupações (CBO), seguida de forma exata nesse nível mais alto de agregação pela base utilizada neste estudo, os microdados da PNAD. Nos demais estados e no próprio PR em 2006, as leis definem pisos para listas de ocupações específicas que nem sempre correspondem à nomenclatura utilizada pela PNAD ou por qualquer outra pesquisa. É o caso dos pisos de SP, analisados neste trabalho após uma compatibilização entre definições feita pelos próprios autores, provavelmente com algumas diferenças em relação às utilizadas em outros estudos. Vale dizer que a maneira como as leis são redigidas dá margem a interpretações variadas não apenas entre os avaliadores de tais políticas, mas também entre empregados, empregadores, sindicatos, fiscais, procuradores e juizes do trabalho, ou até mesmo entre membros dos próprios governos estaduais e das assembleias que as formularam.

Além disso, até 2001, a PNAD adotava outra classificação de ocupações. Para evitar a necessidade de uma segunda compatibilização própria, que poderia reduzir a precisão da análise, este trabalho se concentra no período de 2002 a 2009, explorando, portanto, as oito últimas edições da PNAD da década de 2000. Como a metodologia aqui empregada requer dados do período antes da intervenção, o estudo avalia somente os efeitos de pisos instituídos durante este intervalo: os do PR, criados em 2006, e os de SP, criados em 2007.

Os grupos ocupacionais de cada estado utilizados na análise foram agrupados da seguinte forma. No caso de SP, faz-se uso exatamente dos grupos definidos pelo estado, ou seja, as informações estão desagregadas nos grupos 1, 2 e 3, que correspondem, respectivamente, às ocupações associadas às faixas de pisos I, II e III na legislação paulista, conforme detalha a tabela A.2 do Apêndice A. No caso do PR, para evitar estatísticas baseadas em poucas observações amostrais, optou-se por agregar em apenas três grupos, a partir daqui denominados 5, 6 e 8, as ocupações que a lei paranaense distribuiu por seis faixas de pisos de 2007 a 2009 (tabela A.1 do Apêndice A). O denominado grupo 5 do PR corresponde ao GG 5 da CBO, integralmente coberto pela faixa de piso II desde 2007 até 2009. Por sua vez, os GGs 7 e 8 da CBO, compostos por trabalhadores da indústria cobertos pela mesma faixa de piso V do PR entre 2007 e 2009, foram agregados no que aqui se chama de grupo 8. Finalmente, o grupo 6 também poderia ser chamado de “outros”, pois agrega os GGs 3, 4, 6 e 9 da CBO, que, isoladamente, na amostra da PNAD, têm poucas observações a cada ano, tanto no PR como em outras UFs.<sup>14</sup>

É possível tomar diferentes definições para o cumprimento e o descumprimento de um piso quando se pretende quantificá-los. Para identificar os rendimentos iguais aos pisos, optou-se neste estudo por incluir em tal conjunto, ao redor dos valores exatos de cada piso legal, intervalos que contivessem duas dezenas redondas como limites inferior e superior, pois é nesses valores múltiplos de R\$ 10 que se concentra a maioria dos salários reportados na PNAD. Tendo em vista que os cumpridores do piso podem arredondar seus valores nos contratos de trabalho ou na informação prestada ao entrevistador, aceitou-se que uma observação na pesquisa com renda de R\$ 560, por exemplo, fosse considerada “igual ao piso” da categoria se a lei o tivesse fixado em R\$ 550, em R\$ 569,99 ou em qualquer quantia intermediária. Nas estatísticas a seguir,

<sup>14</sup>Até 2003, a PNAD não cobria as áreas rurais de Acre (AC), Amapá (AP), Amazonas (AM), Pará (PA), Rondônia (RO) e Roraima (RR). Neste trabalho, só foram mantidas na base as observações desses estados coletadas em área urbanizada, mesmo de 2004 a 2009, para que as estatísticas de cada UF se referissem sempre aos mesmos territórios.

este critério evita que as porcentagens de salários iguais aos pisos fiquem altas somente nos pisos com valores exatos múltiplos de R\$ 10, mas, naturalmente, não impede que as frequências permaneçam mais elevadas quando os pisos são iguais ou próximos a múltiplos de R\$ 100, por exemplo.

Mesmo após essa definição, ainda é possível adotar variadas taxas como indicadores de cumprimento. Uma delas é a razão entre o número de pessoas com salário igual ao piso e o total de trabalhadores nas ocupações elegíveis. Outra opção é incluir no numerador todos os trabalhadores com salário igual ou maior que o piso. É possível também mudar o grupo considerado elegível de várias formas, excluindo da conta os trabalhadores com renda menor que o SM ou aqueles com renda muito superior aos pisos, acima de alguma linha arbitrária qualquer. A conta pode ser feita somente para os empregados do setor privado com carteira assinada ou pode incluir outros grupos de ocupados, como empregados sem carteira e trabalhadores por conta própria. A própria renda do trabalho pode estar convertida ou não, pelo número de horas trabalhadas, a um valor proporcional à jornada de 44 horas semanais a que se referem os pisos legais.

A tabela 1 apresenta os valores dos pisos salariais de cada grupo ocupacional definido no PR e em SP. Esses são os valores utilizados (com os devidos ajustes discutidos anteriormente) para marcar quem tem remuneração igual ao piso salarial de seu respectivo estado e grupo ocupacional. Isso vale também para o caso PR, onde se implementou uma agregação dos grupos ocupacionais para fins de análise.

**Tabela 1:** Valores dos pisos salariais legais do PR e SP por ano e mês

Paraná				
Ano	2006	2007	2008	2009
Mês	Junho	Maio	Maio	Maio
Piso I	427,00	462,00	527,00	605,52
Piso II	429,12	464,20	531,00	610,12
Piso III	431,28	466,40	535,00	614,72
Piso IV	433,44	468,60	540,00	620,46
Piso V	435,61	473,00	544,00	625,06
Piso VI	437,80	475,20	548,00	629,65
São Paulo				
Mês	Agosto	Maio	Maio	
Piso I	410,00	450,00	505,00	
Piso II	450,00	475,00	530,00	
Piso III	490,00	505,00	545,00	

Fontes: MTE e assembleias legislativas dos estados.  
Elaboração dos autores.

Optou-se aqui por apresentar dados para empregados do setor privado — com carteira ou sem carteira de trabalho assinada — em ocupações elegíveis para os pisos de cada estado de análise, classificados em faixas de renda associadas ao SM e ao respectivo piso de acordo com suas remunerações mensais do trabalho principal, sem levar em conta o tamanho das jornadas de trabalho.

As tabelas 2 e 3 mostram respectivamente para o PR e SP a distribuição dos trabalhadores da amostra selecionada segundo diferentes faixas de rendimento definidas pelos pisos de cada estado e o salário mínimo nacional. No PR, na média de 2007 a 2009, quando os pisos se referem à totalidade dos grupos analisados, havia aproximadamente 2,6 milhões de empregados - dos quais, 1,8 milhão com carteira - em ocupações com pisos definidos na lei estadual. Esses trabalhadores com piso representavam 91% dos empregados em ocupações bem definidas nos dados da PNAD relativos ao estado. Em SP, de 2007 a 2009, com os pisos referindo-se sempre a ocupações específicas, havia aproximadamente 6,5 milhões de empregados (4,6 milhões com carteira) legalmente cobertos pelos pisos, o equivalente a 52% dos que tinham ocupações bem definidas na base para o estado.

Para anos anteriores à criação dos pisos, foram obtidas taxas simuladas com base nos valores iniciais de cada piso, deflacionados pelo INPC do IBGE e ajustados para as datas de recebimento dos salários informados na PNAD, conforme série disponível no *site* Ipeadata.<sup>15</sup> Nas tabelas 2 e 3, essas taxas simuladas ajudam a extrair uma primeira impressão sobre possíveis efeitos da criação dos pisos sobre as distribuições salariais. Adiante, elas serão úteis para encontrar unidades com distribuições salariais semelhantes ao longo dos anos em que os pisos ainda não existiam.

As tabelas 2 e 3 confirmam que, de modo geral, o contingente com renda *igual ao piso* é pequeno e o com renda *menor que o piso* é grande, o que indica alto grau de descumprimento da legislação estadual, especialmente no PR, onde os valores fixados são mais altos. As porcentagens na faixa considerada igual ao piso oscilam bastante e às vezes se tornam bem mais altos quando os valores legais caem em algum intervalo que inclua uma centena redonda. Parte dessa oscilação ocorre porque essa estatística se baseia em poucas observações da amostra.

As porcentagens com renda menor que o piso apresentam trajetórias mais bem definidas, em geral, caindo logo após a introdução da política e subindo quando os pisos são fortemente aumentados, como no PR em 2008 e 2009. A Figura 1 mostra em quanto as faixas de piso do PR e de SP superam o SM nacional a cada ano, o que ilustra a “*dose do remédio*” ministrado a cada grupo de tratamento, ano a ano. Vale lembrar que as diferenças mostradas na figura, como o aumento no PR em 2008 e 2009, referem-se a níveis também crescentes do próprio SM, que, no mesmo biênio, subiu 4,6% ao ano (a.a.) em termos reais, gerando efeitos em todo o país. Porém, ao avaliar o impacto específico dos pisos introduzidos no PR ou em SP, o que se analisa é o efeito dessas trajetórias exibidas na Figura 1.

---

<sup>15</sup>Disponível em: <http://www.ipeadata.gov.br/>. Vale dizer que, embora tenhamos empregado o INPC geral para o país, as diferenças de variação do índice no período analisado são pequenas entre as regiões metropolitanas (RM) que compõem o INPC. Por exemplo, o INPC brasileiro subiu em média 7% a.a. de 2002 a 2007, enquanto seus componentes da RM de Curitiba e da RM de São Paulo subiram ambos 6,2% a.a. A análise também foi realizada sem deflacionamento e os resultados foram qualitativamente muito semelhantes.

**Tabela 2:** Distribuição dos empregados por faixa de estadual de rendimento no PR (2005-2009) - Em %

		2005 <sup>(1)</sup>	2006 <sup>(2)</sup>	2007	2008	2009
Piso II	Menor que o piso	64,6	55,3	51,8	56,7	64,9
	Menor que o SM	25,2	26,9	24,8	25,2	27,9
	Igual ao SM	19,7	17,6	11,6	13,2	12,3
	Entre o SM e o piso	19,7	10,8	15,4	18,3	24,7
	Igual ao piso	1,4	2,9	2,4	3,3	2,5
	Maior que o piso	34,0	41,8	45,8	40,0	32,7
Pisos I, III, IV e VI	Menor que o piso	30,5	20,9	23,8	26,4	25,6
	Menor que o SM	5,8	5,6	7,8	6,7	6,5
	Igual ao SM	9,4	8,8	5,9	6,6	7,1
	Entre o SM e o piso	15,3	6,5	10,1	13,1	12,0
	Igual ao piso	1,3	1,4	1,2	1,6	7,4
	Maior que o piso	68,2	77,7	75,1	72,0	67,0
Piso V	Menor que o piso	36,4	28,1	28,7	31,6	37,0
	Menor que o SM	6,7	8,4	7,3	5,6	6,0
	Igual ao SM	10,9	10,2	7,6	7,9	6,8
	Entre o SM e o piso	18,8	9,4	13,7	18,2	24,2
	Igual ao piso	1,2	2,5	3,1	3,6	1,5
	Maior que o piso	62,4	69,5	68,2	64,8	61,5
Todos com piso no PR	Menor que o piso	46,3	37,7	37,2	40,3	45,1
	Menor que o SM	14,0	15,5	14,9	13,8	15,0
	Igual ao SM	14,1	13,0	8,9	9,7	9,1
	Entre o SM e o piso	18,2	9,2	13,5	16,8	21,0
	Igual ao piso	1,3	2,4	2,3	2,9	3,5
	Maior que o piso	52,4	59,9	60,4	56,7	51,4

Fonte: PNAD/IBGE. Elaboração dos autores.

Notas: <sup>(1)</sup> Porcentagens de 2005 com base no valor do menor piso inicial, criado em 2006, deflacionado pelo Índice nacional de Preços ao Consumidor (INPC/IBGE). <sup>(2)</sup> Porcentagens para todos os grupos em 2006 com base no menor piso inicial. Em 2006, o PR tinha seis valores de pisos entre R\$ 427,00 e R\$ 437,80, só para ocupações específicas, que representavam 36% dos ocupados no grupo 5; 38% no grupo 6; e 64% no grupo 8. Todas as ocupações dos três grupos passaram a ter piso a partir de 2007.

**Tabela 3:** Distribuição dos empregados por faixa de estadual de rendimento em SP (2006-2009) - Em %

		2006 <sup>(1)</sup>	2007	2008	2009
Piso I	Menor que o piso	41,9	37,5	34,0	32,7
	Menor que o SM	19,6	17,0	19,0	18,3
	Igual ao SM	20,5	12,6	14,0	12,5
	Entre o SM e o piso	1,8	7,8	1,8	1,9
	Igual ao piso	8,2	4,1	6,0	7,6
	Maior que o piso	49,9	58,5	59,3	59,7
Piso II	Menor que o piso	21,3	16,3	13,5	16,4
	Menor que o SM	6,0	5,1	4,9	5,1
	Igual ao SM	7,2	4,8	5,2	5,0
	Entre o SM e o piso	8,0	6,4	3,4	6,3
	Igual ao piso	0,8	2,8	0,9	0,4
	Maior que o piso	77,9	80,9	85,6	83,1
Piso III	Menor que o piso	9,2	8,4	3,8	4,5
	Menor que o SM	2,5	2,3	1,9	1,7
	Igual ao SM	2,3	2,0	0,7	1,2
	Entre o SM e o piso	4,4	4,1	1,2	1,7
	Igual ao piso	0,2	4,3	3,0	1,3
	Maior que o piso	90,6	87,3	93,3	94,2
Todos com piso em SP	Menor que o piso	30,1	25,3	22,3	23,1
	Menor que o SM	12,2	10,3	11,0	11,1
	Igual ao SM	13,2	8,1	8,8	8,2
	Entre o SM e o piso	4,8	6,9	2,5	3,8
	Igual ao piso	4,3	3,5	3,4	4,0
	Maior que o piso	65,6	71,2	74,3	72,9

Fonte: PNAD/IBGE. Elaboração dos autores.

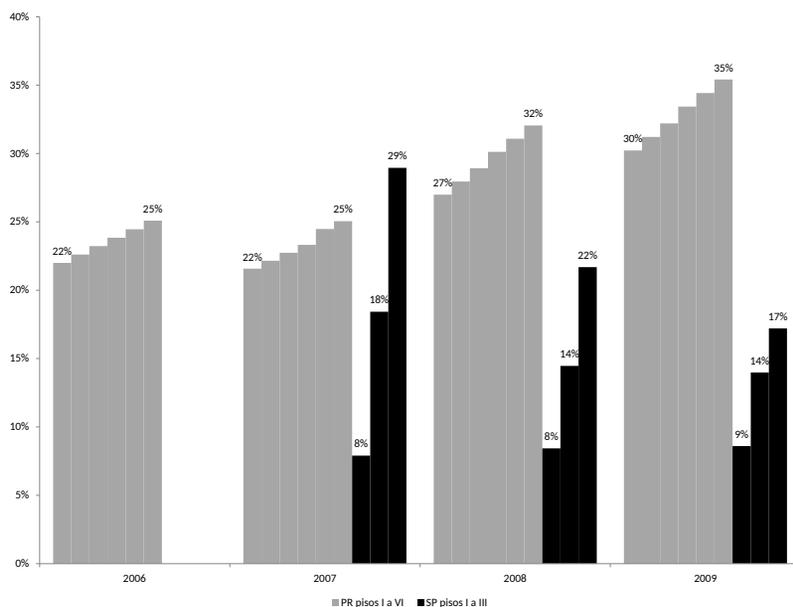
Nota: <sup>(1)</sup> Porcentagens de 2006 com base no valor de cada piso criado em 2007, deflacionado pelo INPC.

## 4 Metodologia

### 4.1 Especificação do modelo

Para computar os efeitos dos pisos estaduais, este trabalho usa o método do grupo de controle sintético inicialmente proposto por Abadie & Gardeazabal (2003) e posteriormente formalizado por Abadie et al. (2010). O método baseia-se na construção de um grupo de comparação que seja o mais semelhante - em um sentido explicitado adiante - ao grupo de tratamento antes da intervenção. A ideia básica é que, se o grupo de controle sintético tem comportamento similar ao da unidade tratada até o início da intervenção, então é de se esperar que as diferenças nas variáveis de resultado que porventura apareçam entre os dois grupos após o programa representem o efeito de interesse.

Uma das principais vantagens do método é que ele procura combinar as unidades que compõem o grupo de controle sintético de forma flexível. Utilizando informações sobre a própria variável de resultado e um conjunto de preditores dessa variável, o método busca um vetor de pesos para as potenciais unidades formadoras do grupo de controle, de tal forma que o grupo



Fonte: Assembleias legislativas dos estados. Elaboração dos autores.

Nota: As alturas das barras para cada estado correspondem às diferenças dos pisos dos respectivos grupos referenciados em algarismos romanos de forma crescente.

**Figura 1:** Pisos salariais de PR e SP: diferenças sobre o SM nas datas das PNADs (2006-2009)

sintético final seja bem próximo às características da unidade tratada previamente à intervenção. Outros métodos que utilizam um grupo de controle construído a partir de potenciais unidades de comparação, como o diferenças em diferenças, tendem a ser não somente menos flexíveis, mas também menos transparentes, uma vez que são frequentemente baseados em percepções subjetivas dos analistas sobre as semelhanças entre os grupos de tratamento e comparação. Além disso, a escolha de uma unidade como grupo de controle nem sempre é feita sem que o analista já tenha, antes de realizar a estimação, informações sobre a evolução de seus indicadores no período da intervenção. Nesse sentido, um grupo de controle sintético é também composto de forma mais “cega”, a partir de critérios objetivos.<sup>16</sup>

A fim de formalizar o funcionamento do método, suponha que existem informações sobre  $J + 1$  unidades (no caso, UFs) e que apenas uma delas seja a unidade tratada (SP ou PR). As demais  $J$  unidades, portanto, formarão um conjunto de unidades que poderão compor o grupo de controle sintético (esse conjunto é denominado de conjunto de unidades doadoras). Assuma que as  $J$  unidades doadoras não tenham sido nem direta nem indiretamente afetadas pela intervenção. Na prática, é difícil verificar se as potenciais unidades de controle não foram de fato afetadas indiretamente pela intervenção. Quanto a isso, assim como em outros métodos, cabe ao analista ser criterioso na seleção do conjunto de unidades de comparação.

<sup>16</sup>O método do grupo de controle sintético é particularmente apropriado para análises de efeitos de intervenções em unidades agregadas, como é o caso da política de pisos salariais por grupos ocupacionais das UFs em foco neste artigo. Uma das vantagens do uso de dados agregados é que ele tende a minorar erros de medida tipicamente presentes em dados no nível do indivíduo.

Suponha que nos dados existam  $t = 1, \dots, T$  períodos de observação, sendo  $t = 1, \dots, T_0$  o intervalo pré-tratamento,<sup>17</sup> com  $T_0 < T$ . No caso do PR, por exemplo,  $t = 2002, \dots, 2009$  com  $T_0 = 2005$ .

Seja  $Y_{it}^I$  a variável de resultado da unidade  $i$  no período  $t$  se esta unidade for tratada e  $Y_{it}^N$  a mesma variável de resultado para a unidade  $i$  no período  $t$  caso esta unidade não fosse tratada. Como  $Y_{it}^N$  representa o contrafactual da variável de resultado, é possível definir o efeito do programa para a unidade  $i$  no tempo  $t$  como:  $a_{it} = Y_{it}^I - Y_{it}^N$ . Assim, utilizando a notação de resultados potenciais, o analista observa para a unidade  $i$  no período  $t$ :  $Y_{it} = Y_{it}^N + a_{it}D_{it}$ , onde  $D_{it}$  é uma variável que assume valor unitário se a unidade  $i$  for exposta à intervenção em  $t$  e valor nulo caso contrário. Sem perda de generalidade, considerando  $i = 1$  como a unidade tratada (SP ou PR neste trabalho), o objetivo é estimar a sequência  $\{a_{1T_0+1}, \dots, a_{1T}\}$ .

Dado que  $Y_{1t}^I$  é observado diretamente dos dados, resta encontrar um modo de estimar o contrafactual  $Y_{1t}^N$ . Abadie et al. (2010) propõem o seguinte processo que gera os valores desta variável:

$$Y_{1t}^N = \delta_t + \theta_t Z_i + \pi_t \mu_i + \varepsilon_{it}, \quad (1)$$

em que  $\delta_t$ , denota um fator comum a todas as unidades no período  $t$ ,  $Z_i$  representa um vetor ( $r \times 1$ ) de características preditoras da variável de resultado,  $\theta_t$  é um vetor ( $1 \times r$ ) de parâmetros,  $\pi_t$  é um vetor ( $1 \times F$ ) de fatores comuns que entra no processo gerador multiplicando o vetor ( $F \times 1$ ) de *factor loadings*  $\mu_i$  e  $\varepsilon_{it}$  é um choque aleatório com média zero.<sup>18</sup>

O próximo passo é encontrar o vetor de pesos para combinar as  $J$  unidades doadoras em um grupo de controle sintético. Seja  $W = (w_2, \dots, w_{J+1})$  um vetor de pesos tal que  $0 \leq w_j \leq 1$ ,  $j = 2, \dots, J$  e  $\sum_{j=2}^{J+1} w_j = 1$ . Cada vetor distinto  $W$  corresponde a um grupo de controle sintético potencial, de forma que, com base nos dados observados para  $Y$  e  $Z$  das unidades doadoras e no processo que gera a variável de resultado proposto em 1, é possível escrever:

$$\sum_{j=2}^{J+1} w_j Y_{jt} = \delta_t + \theta_t \sum_{j=2}^{J+1} w_j Z_j + \pi_t \sum_{j=2}^{J+1} w_j \mu_j + \sum_{j=2}^{J+1} w_j \varepsilon_{jt}. \quad (2)$$

Suponha que exista um vetor de pesos  $W^* = (w_2^*, \dots, w_{j+1}^*)$  tal que:

$$\sum_{j=2}^{J+1} w_j^* Y_{j1} = Y_{11}, \dots, \sum_{j=2}^{J+1} w_j^* Y_{jT_0} = Y_{1T_0} \text{ e } \sum_{j=2}^{J+1} w_j^* Z_{j1} = 1. \quad (3)$$

<sup>17</sup> Assume-se que a intervenção não tenha impactos nas variáveis de resultado antes de  $T_{0+1}$ , via efeitos antecipatórios por exemplo. Note-se que, quando existam efeitos desse tipo, o período  $T_0$  pode ser incluído no intervalo em que o efeito do programa já está potencialmente operando.

<sup>18</sup> Note que, ao permitir que o efeito de características não observadas da unidade  $i$  varie no tempo, esse modelo torna-se mais geral que o tradicionalmente empregado na abordagem de diferenças-em-diferenças. Neste último,  $\pi_t$  assumiria um valor constante para todos os períodos de tempo, o que permitiria eliminar o fator não-observável  $\mu_i$  ao se computar a diferença da equação 1 entre períodos de tempo.

Abadie et al. (2010) mostram que, sob certas condições de regularidade padrão,<sup>19</sup> a diferença

$$Y_{1t}^N - \sum_{j=2}^{J+1} w_j^* Y_{jt}.$$

será próxima de zero, ou seja, o vetor  $W^*$  permite criar uma média ponderada da variável de resultado observada nas unidades doadoras cujo valor é muito próximo do contrafactual da unidade tratada para o período  $t$ . Mas isso é o que geralmente se busca nos métodos não experimentais de avaliação de impactos de programas. Assim, utilizando essa média ponderada, tem-se uma estimativa não-viesada do impacto de interesse no período  $t$ :  $\widehat{a}_{it} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$ .

O último passo do método é encontrar o vetor  $W^*$ .<sup>20</sup>

Seja  $X_1$  uma matriz ( $k \times 1$ ) que contém os valores das características  $Z_1$  e da variável de resultado (e/ou combinações lineares dela) para a unidade tratada no intervalo pré-intervenção. Por exemplo,

$$X_1 = \left( Z_1, Y_{11}, \dots, Y_{1T_0}, T_0^{-1} \sum_{t=1}^{T_0} Y_{1t} \right).$$

Denotando por  $X_0$  a matriz  $k \times j$  com as mesmas variáveis de  $X_1$  para as  $j$  unidades doadoras, o método procura minimizar a distância  $\|X_1 - X_0 W\|$ . Para essa minimização, Abadie et al. (2010) propõem utilizar uma matriz  $V$ , ( $k \times k$ ) simétrica e positiva semidefinida, que mede a distância de interesse por meio da fórmula:

$$\|X_1 - X_0 W\|_V = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)}. \quad (4)$$

A matriz  $V$  será, portanto, uma matriz que estará (implicitamente) dando peso às variáveis que compõem a matriz de diferença  $X_1 - X_0 W$ .

Notando que a escolha de  $V$  influencia o erro quadrático médio (EQM) do estimador proposto, Abadie et al. (2010) seguem o procedimento já adotado em Abadie & Gardeazabal (2003), a saber:  $V$  foi escolhida entre as matrizes diagonais e positivas definidas de forma a minimizar o EQM predito da variável de resultado durante o período pré-intervenção. Este artigo segue o mesmo procedimento.

## 4.2 Operacionalização

Para estimar os efeitos dos pisos regionais, delimitou-se a amostra para incluir somente os trabalhadores do setor privado empregados com ou sem carteira.

<sup>19</sup>As principais condições de regularidade são que a matriz  $\sum_{t=1}^{T_0} \pi_t' \pi_t$  seja não-singular e que o número de períodos de tempo pré-intervenção seja elevado relativamente à escala dos choques aleatórios  $\varepsilon_{it}$ .

<sup>20</sup>Na prática, pode ocorrer de não existir um vetor de pesos que faça a diferença  $Y_{it}^N - \sum_{j=2}^{J+1} w_j^* Y_{jt}$  ser próxima de zero. Isso ocorrerá quando  $\{(Y_{11}, \dots, Y_{1T_0}, Z_1')\}$  não pertencer ao conjunto convexo formado por  $\{(Y_{21}, \dots, Y_{2T_0}, Z_2'), \dots, (Y_{j+11}, \dots, Y_{j+1T_0}, Z_{j+1}')\}$ .

Adicionalmente, restringiu-se a amostra para o conjunto de trabalhadores cujas ocupações estejam em um dos grupos ocupacionais para os quais há um piso salarial definido nas respectivas legislações de SP e PR (ver seção 3). Essa delimitação se deve ao fato de que, pelo menos em princípio, esses são os trabalhadores mais diretamente afetados pela intervenção.<sup>21</sup>

O conjunto de unidades doadoras foi definindo com base na exclusão das UFs que possuem uma política de piso salarial ocupacional. Assim, como até 2009 havia quatro UFs que tinham essa política estabelecida (RJ, RS, PR e SP), o conjunto de UFs doadoras para SP ou PR foi composto pelas 23 demais UFs do país.<sup>22</sup> Quando o nível de análise passa a ser o grupo ocupacional, o conjunto de unidades doadoras foi formado por até três grupos ocupacionais de cada uma destas 23 UFs. Foram excluídos apenas, do total de 69 grupos disponíveis, aqueles que tivessem, em algum ano, menos de trinta observações de interesse na base para compor o numerador dos indicadores analisados ou taxas de 100% nesses indicadores. Apresentamos na tabela 3 do Apêndice A uma síntese de estatísticas referentes a frequência amostral de cada UF.<sup>23</sup>

Foram estimados efeitos dos pisos do PR e de SP nas seguintes variáveis de resultado (Y):

- porcentagem situada abaixo dos pisos analisados: proporção dos empregados cujo salário no trabalho principal se encontrava estritamente abaixo do valor do piso do grupo de tratamento analisado;<sup>24</sup>
- taxa de emprego: razão entre o total de empregados com ou sem carteira no grupo e o total de indivíduos na população em idade ativa (PIA) da UF, definida como população com 10 anos de idade ou mais;
- grau de informalidade: proporção dos sem carteira no total de empregados com renda de até 1,5 SM;<sup>25</sup> e
- hiato salarial formal-informal: logaritmo natural da razão entre a média salarial dos empregados com carteira e a dos sem carteira.

Vale observar que a porcentagem situada abaixo do piso não representa uma variável de resultado típica. No entanto, a análise do “feito” dos pisos

<sup>21</sup> Como mencionado na seção 3, esse grupo é amplo, representando 91% e 52% dos empregados no PR e em SP, respectivamente. O método empregado neste artigo procura encontrar um grupo de controle sintético com base em informações agregadas dos trabalhadores que se encontram nas mesmas ocupações das demais UFs doadoras. Embora nada garanta que os trabalhadores dessas ocupações sejam parecidos entre as UFs em características não observáveis, o método implicitamente controla para diferenças agregadas entre a UF de tratamento e as UFs doadoras. Naturalmente, há sempre um problema de validade externa, ou seja, não é possível garantir que os resultados obtidos para a UF de tratamento possam ser estendidos para (os mesmos grupos ocupacionais) de UFs não tratadas.

<sup>22</sup> Como só introduziu uma política de pisos salariais em 2010, SC pertence ao conjunto de UFs doadoras.

<sup>23</sup> Tabelas com o número de observações anuais em cada grupo de cada UF utilizada nas estimações podem ser requisitadas aos autores.

<sup>24</sup> Ao estudar o efeito sobre esse indicador no grupo 6 do PR, utiliza-se o menor dos pisos deste grupo (faixa I) como referência para todas as observações da base. Quando a unidade tratada em análise é uma UF agregada (PR ou SP), a referência para cada observação é o piso associado ao seu respectivo grupo ocupacional na UF tratada.

<sup>25</sup> Em princípio, os trabalhadores por conta-própria não têm seus rendimentos vinculados nem aos valores estabelecidos para os pisos estaduais nem ao salário mínimo. Por essa razão, esse indicador utiliza apenas assalariados sem carteira assinada para medir o impacto de interesse sobre a informalidade.

salariais sobre ela é importante, uma vez que o cumprimento ou descumprimento da lei pode ser interpretado como o principal canal pelo qual a política afetará outras variáveis do mercado de trabalho. Outra observação sobre a escolha das variáveis é que não foi incluída nem a taxa de desemprego nem a taxa de participação. A exclusão se deve ao fato de que a análise empírica foi realizada somente para os trabalhadores que se encontravam nas categorias ocupacionais cobertas pelas respectivas legislações estaduais dos pisos salariais de SP ou PR. Por isso, não seria possível demarcar com os dados disponíveis quem é desocupado para certas ocupações.

No modelo proposto por Abadie et al. (2010), as matrizes  $(X_1, X_0)$  são formadas pelas variáveis preditoras em  $Z$  e a própria variável  $Y$  (e/ou combinações lineares de seus valores) no período pré-programa. Os resultados aqui obtidos foram baseados na construção de  $(X_1, X_0)$  que envolveu, além da variável  $Y$  em cada um dos anos anteriores ao início da política em SP e no PR, a seguinte lista selecionada com seis possíveis preditores dos indicadores de resultado: nível médio de escolaridade (em anos de estudo), proporção de empregados com até 24 anos de idade, proporção de empregados na indústria, proporção de empregados homens, proporção de empregados que declaram ter cor branca, renda domiciliar *per capita* na UF.

Mais especificamente, foi utilizado o valor médio por unidade (UF ou grupo ocupacional, conforme o caso) dessas seis covariadas no período desde 2002 até o ano anterior à introdução da lei dos pisos salariais em SP (2006) e no PR (2005). Entretanto, em todos os exercícios, a rotina de minimização do EQM predito obteve as melhores aproximações entre unidades tratadas e seus respectivos grupos sintéticos atribuindo peso próximo a zero para todos os seis preditores; e distribuindo praticamente todo o peso entre os resultados anuais pré-tratamento da variável de resultado  $Y$ .

A título de exemplo, as tabelas A.3 e A.4 do Apêndice A mostram, respectivamente, a matriz  $V$  e a matriz  $W$  que levam ao resultado do exercício reportado na subseção 5.1.<sup>26</sup> A tabela A.3 do Apêndice A exibe também os valores pré-tratamento da variável *porcentagem abaixo do piso*( $Y$ ) e das médias das seis covariadas (matriz  $Z$ ) observados na unidade tratada e em seu sintético construído. Nesse exemplo, como as covariadas recebem pequeno peso na minimização do EQM, é possível construir um sintético que não se parece com a unidade tratada em termos de alguns dos preditores. Quanto à proporção de pessoas brancas, em particular, a diferença é bem grande. Por sua vez, a unidade tratada e seu sintético se aproximam muito bem em termos da variável  $Y$  em todos os anos do período pré-tratamento. Vale observar que os pesos finais das UFs que compõem o grupo de controle sintético são internos ao método, ou seja, eles não necessariamente levam em consideração dimensões como a proximidade espacial entre elas. Por exemplo, como mostra a tabela A.4 no Apêndice A, o Amazonas recebe um peso considerável na construção do grupo de controle do Paraná para o indicador da proporção que recebe abaixo do piso.

### 4.3 Inferência

A fim de fazer inferência sobre as estimativas obtidas dos efeitos de interesse, foi adotado também o método proposto em Abadie et al. (2010). Esse método

---

<sup>26</sup>As matrizes  $V$  e  $W$  dos demais exercícios realizados podem ser requisitadas aos autores.

é inspirado no teste de permutação, que é comumente empregado na área de estatística e vem sendo cada vez mais aplicado em econometria.<sup>27</sup> A ideia é comparar o efeito estimado para a unidade efetivamente tratada com estimativas do efeito em estudo, supondo-se que as unidades doadoras tivessem sido tratadas. Sob a hipótese de que a intervenção não afetou outras unidades além da unidade efetivamente tratada, a comparação das estimativas obtidas permite avaliar a distância entre a estimativa do efeito para a unidade tratada e as estimativas calculadas para as demais unidades. Quanto mais distante estiver a estimativa para o grupo tratado relativamente às demais estimativas, mais forte é a evidência de que o efeito do tratamento não ocorreu por acaso. Uma das vantagens desse método é que ele não demanda um número elevado de unidades pertencentes ao conjunto de unidades doadoras.

A operacionalização desse método de inferência foi realizada estimando-se o modelo da subseção 4.1 separadamente para cada UF pertencente ao conjunto de unidades doadoras de SP ou PR. Especificamente, estimou-se para cada uma das 23 UFs desse conjunto a diferença entre os valores da variável de resultado para a UF e seu grupo sintético falseado. Durante o período pós-intervenção, essa diferença de valores pode ser denominada efeito placebo. Procedeu-se de forma semelhante para as 58 unidades doadoras do grupo 1 de SP e para as 68 unidades doadoras do grupo 5 do PR.

## 5 Resultados

Conforme explicado na introdução, foram selecionados quatro indicadores para mapear em que medida os pisos salariais estaduais afetam o desempenho do mercado de trabalho: proporção de empregados com remuneração inferior ao valor do piso; taxa de emprego; grau de informalidade; e hiato salarial entre empregados com carteira e sem carteira. As subseções 5.1 e 5.2 analisam os resultados relativos aos dois primeiros e a 5.3 analisa os dois últimos.

### 5.1 Proporção de empregados com remuneração inferior ao piso

Os resultados foram obtidos mediante o emprego do método discutido na seção 4 em diversas agregações alternativas dos dados, tanto para PR como para SP. Para cada UF, os resultados serão reportados na forma gráfica tanto para o agregado das ocupações com piso legislado como para cada um de três grupos ocupacionais com diferentes valores de piso.

A primeira Figura será discutida em mais detalhes para que o leitor se acostume com a forma de apresentação das informações adotada até o fim do texto. A Figura 2 reporta quatro linhas referentes ao agregado das ocupações com piso no PR. As duas linhas intermediárias são de fundamental importância para a compreensão do método. A linha cinza contínua mostra a evolução de fato da porcentagem de empregados com remuneração abaixo do respectivo piso no PR entre 2002 e 2009. A linha preta tracejada mostra a evolução deste mesmo indicador observada no grupo sintético que procura replicar a evolução no PR de fato durante o período anterior à criação do piso.

<sup>27</sup>O teste de permutação é tipicamente realizado com base na construção da distribuição de uma estatística de teste (uma estatística  $z$ ). Essa distribuição é obtida a partir da permutação aleatória das unidades da amostra entre os grupos de tratamento e de controle. Ver Angrist & Krueger (1999) para uma discussão sobre o uso desse tipo de teste na área de economia aplicada.

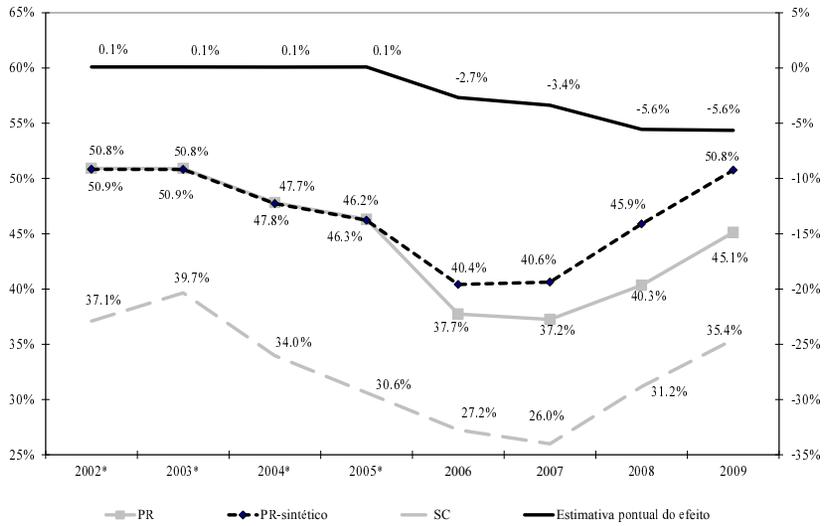
A linha do PR de fato (cinza contínua) mostra um claro declínio da parcela de trabalhadores com remuneração abaixo do piso entre 2005 (46,3%) e 2007 (37,2%), quando a política cobre todas as ocupações analisadas. Logo, uma simples comparação da evolução temporal desse indicador poderia levar à suposição de que a política tivesse feito diminuir em 9 pontos percentuais (p.p.) a proporção de trabalhadores com remuneração inferior ao piso. No entanto, a distribuição de rendimentos do PR pode ter sido afetada por outros motivos, alguns dos quais inclusive podem ter contribuído para diminuir a parcela de trabalhadores com renda menor que o piso. A evolução do grupo sintético tenta mostrar qual seria essa evolução quando afetada pelos demais fatores e sem a influência do piso. Segundo o método empregado, uma considerável parcela da queda desse indicador, quase 6 p.p., ocorreria mesmo sem a introdução do piso no PR. A estimativa do impacto do piso deve desconsiderar, portanto, essa parcela da queda que não pode ser atribuída ao piso.

A estimativa do efeito do piso consiste na diferença entre o PR e seu sintético, representada pela linha preta contínua, com valores orientados pelo eixo vertical da direita. Essa estimativa aponta que a introdução do piso induziu a uma queda de 2,7 p.p. na parcela de trabalhadores com remuneração inferior ao piso em 2006 e de 3,4 p.p. em 2007. Vale notar que o ajuste da evolução do grupo sintético à observada no PR no período pré-tratamento é quase perfeito. Isso dá suporte à interpretação de que as quedas de 2006 e 2007 se devem efetivamente à introdução do piso.

A linha cinza tracejada mostra a evolução em SC, inserida nesse gráfico para contrastar o método empregado com o diferenças em diferenças. Nas aplicações desse método para estimativas de efeito de pisos salariais estaduais, é praxe o uso de algum estado vizinho que não teve alteração no valor do piso como grupo de controle, com sua evolução no papel desempenhado aqui pelo grupo sintético. Em qualquer um dos casos, o grupo de controle deveria ser capaz de replicar a evolução do grupo de tratamento, mesmo que em patamar distinto, antes da criação do piso. A Figura 2 mostra que a coincidência ocorre quando se usa o grupo sintético como controle. A evolução reportada para esse grupo, como já mencionado, replica quase exatamente a evolução observada para o PR. Tanto nesse grupo de controle como no PR de fato, a porcentagem de trabalhadores recebendo uma quantia inferior ao piso fica estável entre 2002 e 2003, cai cerca de 3 p.p. de 2003 a 2004 e mais 1,5 p.p. entre 2004 e 2005. Por sua vez, a evolução observada em SC difere bastante desse padrão. Nesse estado, a porcentagem de trabalhadores com remuneração abaixo do piso paranaense sobe 2,5 p.p. entre 2002 e 2003 para em seguida cair 5,7 p.p. em 2004 e 3,4 p.p. em 2005.

Em comum com a Figura 2, os próximos gráficos trarão somente a linha preta contínua referente ao efeito do piso salarial, acompanhada de outras linhas que fazem parte do procedimento de inferência comentado na subseção 4.3. A Figura 3 dá sequência à análise da porcentagem de trabalhadores com remuneração abaixo do piso no PR e é a primeira de uma série com esse formato. As linhas finas e claras são usadas para inferência e representam diferenças análogas entre a evolução de cada uma das unidades não tratadas e seu respectivo controle sintético. Se a linha grossa for uma das mais distantes do zero em qualquer uma das direções, para cima ou para baixo, então se pode inferir que o grupo tratado foi realmente afetado.

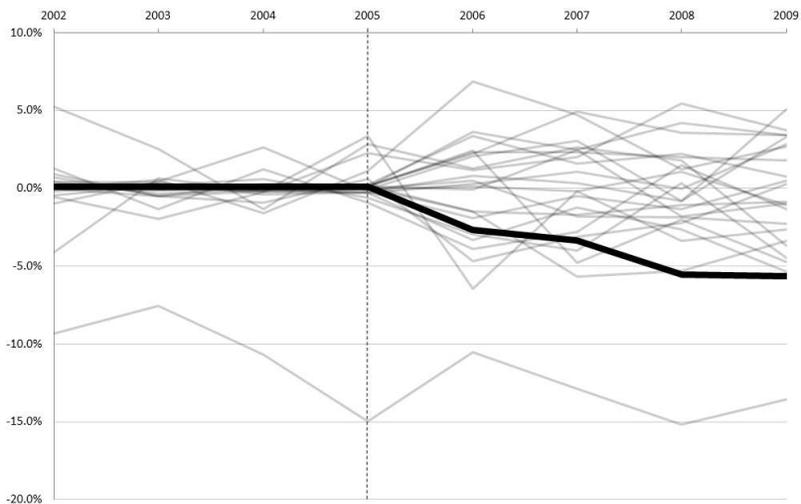
No PR, a qualidade do ajuste do contrafactual é muito boa. Isto pode ser comprovado pela comparação no período pré-intervenção entre a linha



Fonte: PNAD/IBGE Elaboração dos autores.

\* De 2002 a 2006, porcentagens com base no menor piso do PR em 2006, deflacionado pelo INPC até 2005.

**Figura 2:** Empregados com rendimento abaixo do piso e efeito da lei sobre esse indicador - Paraná\*

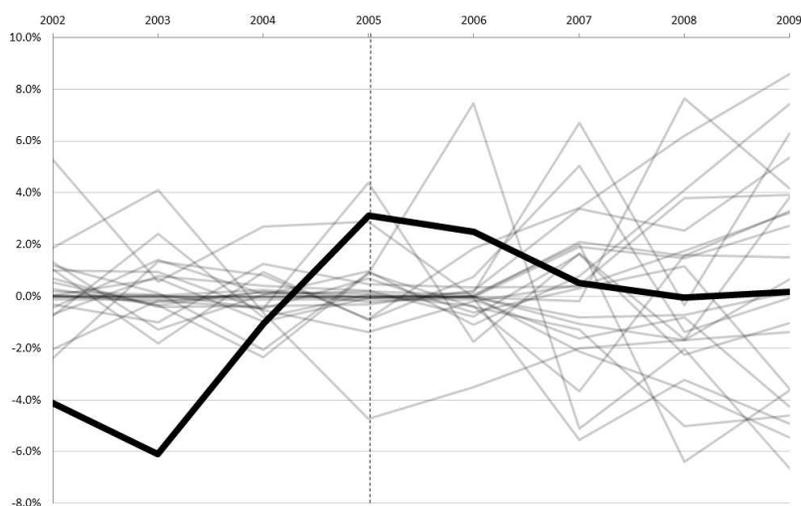


Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 3:** Abaixo do piso: PR e demais UF's

cheia e a que representa o zero, que aparecem sobrepostas. A trajetória pós-tratamento aponta a existência de algum efeito a partir de 2006, crescendo até se tornar significativo em 2008 e 2009, quando a linha do PR passa a ser a mais negativa entre as UFs que se ajustaram bem aos seus respectivos sintéticos nos anos pré-tratamento. Na média de 2007 a 2009, período em que a legislação cobre todas as categorias estudadas, estima-se que 4,9% dos empregados nas ocupações com piso no PR deixaram de ganhar menos que seus respectivos pisos devido à introdução da política. Isto equivale a 10,6% dos empregados com ou sem carteira que supostamente ganhariam menos que o valor do piso se ele não existisse. Os resultados são similares - embora, curiosamente, um pouco menos significativos do ponto de vista estatístico - quando o exercício é reproduzido para uma amostra restrita somente aos empregados com carteira assinada.

A Figura 4 é feita de forma análoga para SP. Nessa UF o ajuste é bem precário, indicando que não foi possível agregar os trabalhadores das demais UFs sem pisos estaduais de forma a replicar a evolução em SP da proporção de empregados com remuneração abaixo do piso. A mesma conclusão vale para a amostra restrita aos empregados com carteira.



Fonte: PNAD/IBGE Elaboração dos autores.

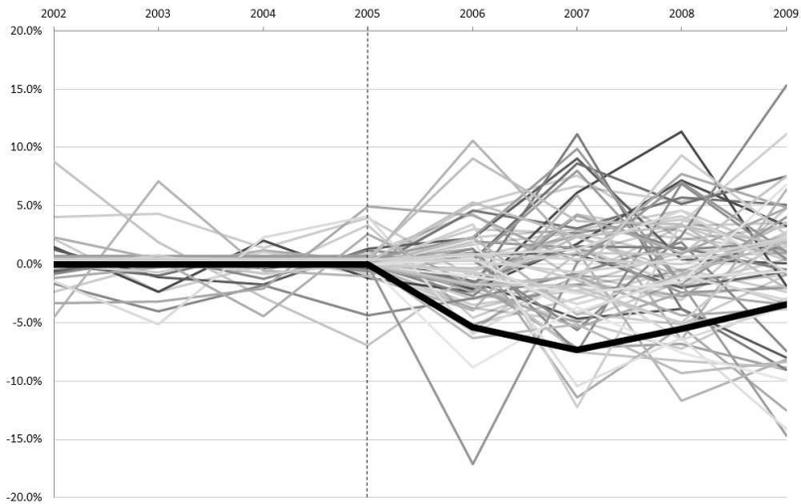
**Figura 4:** Abaixo do piso: SP e demais UFs

O resultado — ou a falta dele — apontado para SP motiva a investigar cada um dos grupos ocupacionais de forma isolada. Mesmo no caso do PR, o aparente baixo efeito reportado pode ser fruto da agregação dos grupos ocupacionais.

Os próximos seis gráficos comprovam que há uma heterogeneidade do efeito entre os grupos e que ao menos um dos grupos ocupacionais de cada UF de fato teve a proporção de empregados com remuneração abaixo do piso afetada pelo piso. No caso do PR, as Figuras 5 a 7 mostram, em um extremo, o agregado de ocupações que corresponde ao GG 5 da CBO sendo claramente afetado. A Figura 5 indica que essa proporção caiu cerca de 8 p.p. dois anos após a introdução do piso. Vale lembrar que, em 2006, nem todas as ocupações do GG 5 da CBO estavam contempladas na legislação do piso paranaense,

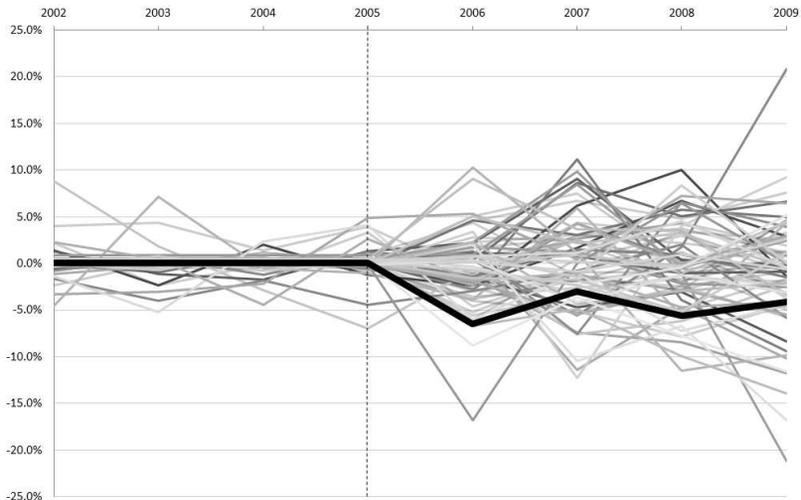
mas em 2007 todas passaram a ser contempladas. A estimativa é que, na média de 2007 a 2009, 5,4% de todos os empregados do grupo 5 do PR tenham deixado de ganhar menos que seu piso por causa da legislação estadual. Essa porcentagem corresponde a 8,6% do contingente que supostamente receberia menos que o valor do piso caso a lei não existisse.

No outro extremo do caso paranaense, está o resultado para o grupo ocupacional 8 na Figura 7, no qual uma considerável parte das linhas cinzas abaixo da linha preta indica uma forte possibilidade de não haver efeitos significativos. A Figura 6 exhibe resultados do grupo 6, para o qual, se há algum efeito, este parece ser marginal e restrito ao ano de 2006.



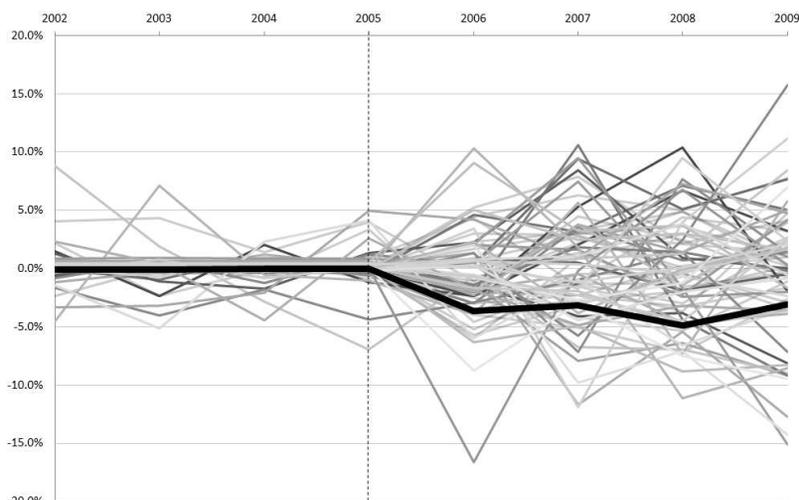
Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 5:** Abaixo do piso: grupo 5 do PR e grupos das demais UFs



Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 6:** Abaixo do piso: grupo 6 do PR e grupos das demais UFs



Fonte: PNAD/IBGE Elaboração dos autores.

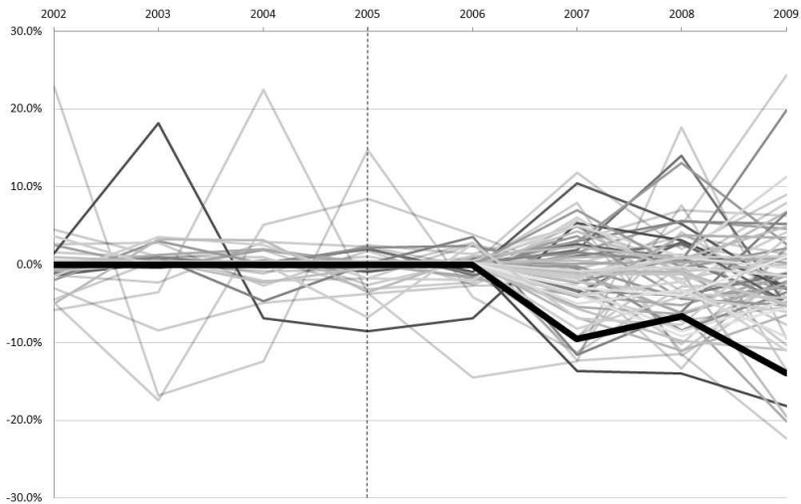
**Figura 7:** Abaixo do piso: grupo 8 do PR e grupos das demais UF's

As Figuras 8 a 10 indicam que os grupos ocupacionais 2 e 3 têm efeitos não significativos, enquanto o grupo 1 apresenta um efeito significativo em 2007 e 2009, porém apenas marginalmente em 2008. A Figura 8 mostra que a proporção de trabalhadores do grupo 1 com remuneração abaixo do piso cai em relação à evolução do grupo de controle sintético. A diferença chega a 9,5 p.p. no primeiro ano de vigência do piso, 6,7 p.p. no segundo e 13,9 p.p. no terceiro. Estimou-se que, na média dos três anos analisados de vigência do piso, 10% dos empregados do grupo 1 de SP passaram a ganhar pelo menos tanto quanto seu piso devido à lei estadual. Em contraste ao caso de SP-agregado (Figura 4), o ajuste do grupo de controle sintético à evolução observada no grupo 1 antes do piso é muito bom.

Na Figura 9, o grupo 2 também apresenta um bom ajuste. Entretanto, esse grupo não parece ter sido afetado pelo piso. A Figura 10 mostra que o grupo 3 não teve ajuste suficiente para oferecer um resultado muito crível. Esse desajuste, no grupo mais bem remunerado de uma UF com salários relativamente altos, evidencia a dificuldade em se obterem grupos de controle adequados para grupos de tratamento muito próximos aos extremos, uma restrição válida também para outros métodos, embora nem sempre evidente em seus resultados.

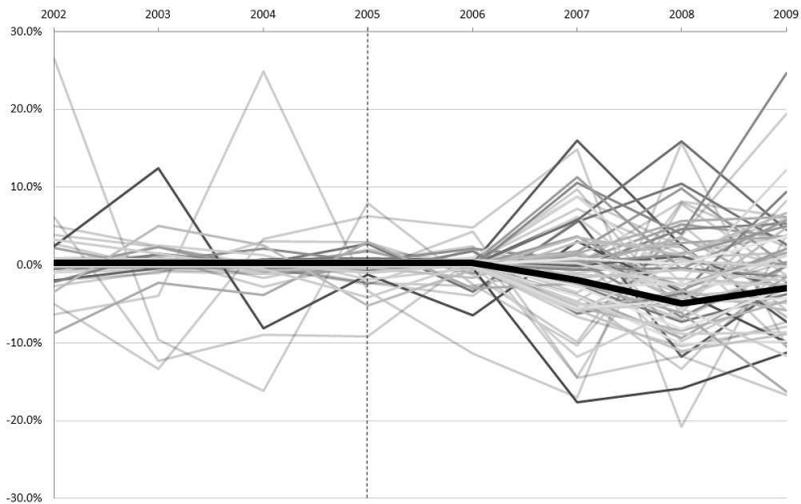
Para confirmar se os efeitos reportados para a proporção de empregados com remuneração abaixo do piso, tanto no grupo 5 do PR como no grupo 1 de SP, são realmente provenientes da introdução dos respectivos pisos, foi conduzido um teste de falseamento. Nesse teste, simula-se que a política tenha sido introduzida um ano antes, ou seja, repetiram-se os respectivos exercícios de estimação e inferência como se o piso tivesse sido introduzido no PR em 2005 e em SP em 2006.

Se as Figuras apresentarem qualquer distanciamento nesses anos de tratamento placebo, é provável que o efeito capturado seja proveniente de outro fator, pois ainda não havia pisos legislados nesses anos. No caso do PR, o método parece ser capaz de isolar o efeito do piso de outros efeitos. A Figura 11 mostra que as linhas para a evolução observada do grupo 5 e do seu



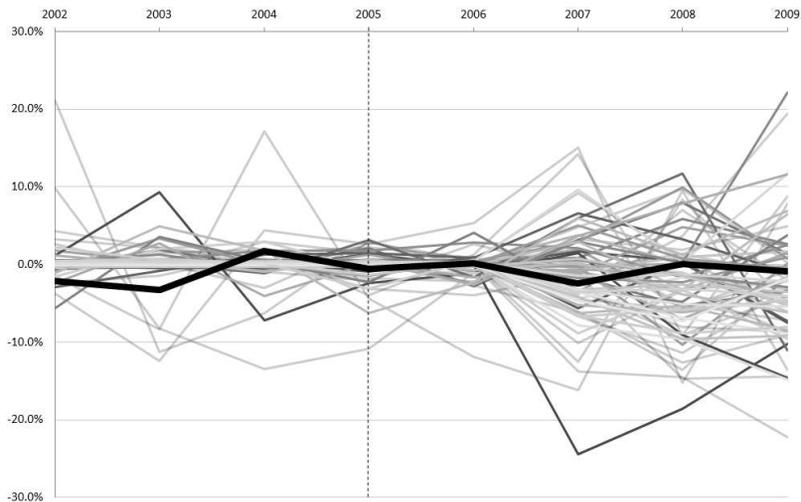
Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 8:** Abaixo do piso: grupo 1 de SP e grupos das demais UFs



Fonte: PNAD/IBGE Elaboração dos autores.

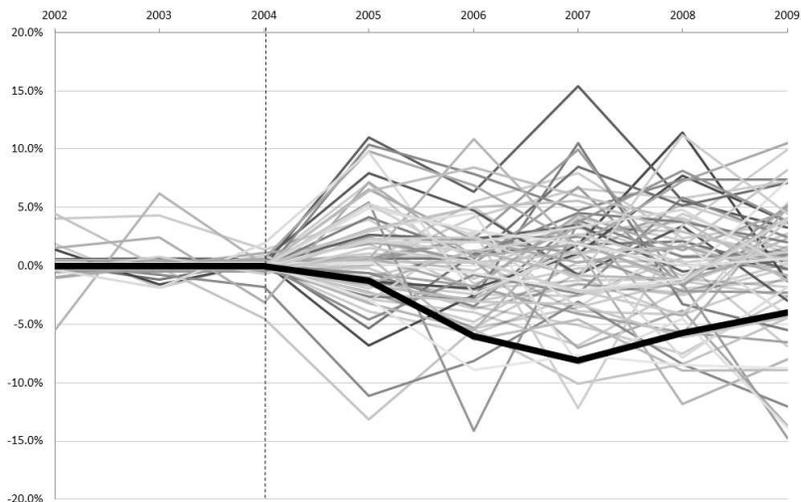
**Figura 9:** Abaixo do piso: grupo 2 de SP e grupos das demais UFs



Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 10:** Abaixo do piso: grupo 3 de SP e grupos das demais UF's

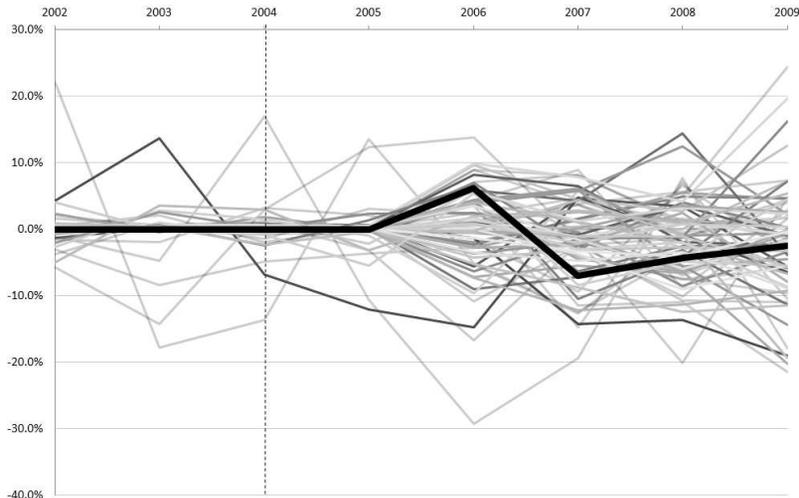
contrafactual continuam sobrepostas em 2005. Essa evidência reforça que o efeito reportado na Figura 5 pode ser atribuído aos respectivos pisos estaduais. No caso de SP, a Figura 12 aponta que o método talvez não consiga isolar de forma perfeita o efeito do piso, haja vista um distanciamento do zero já em 2006, antes da criação do piso. Isto significa que incluir ou não os dados de 2006 como critério de aproximação leva a grupos sintéticos com trajetórias diferentes. Esse fato suscita dúvidas sobre a capacidade de identificar um efeito realmente associado ao piso na Figura 8.



Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 11:** Placebo em  $t - 1$  para abaixo do piso no grupo 5 do PR

Em suma, a análise do primeiro indicador apontou que apenas no grupo ocupacional 5 do PR e no grupo 1 de SP parece ter havido efeitos significativos dos respectivos pisos salariais sobre a proporção de trabalhadores com



Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 12:** Placebo em  $t - 1$  para abaixo do piso no grupo 1 de SP

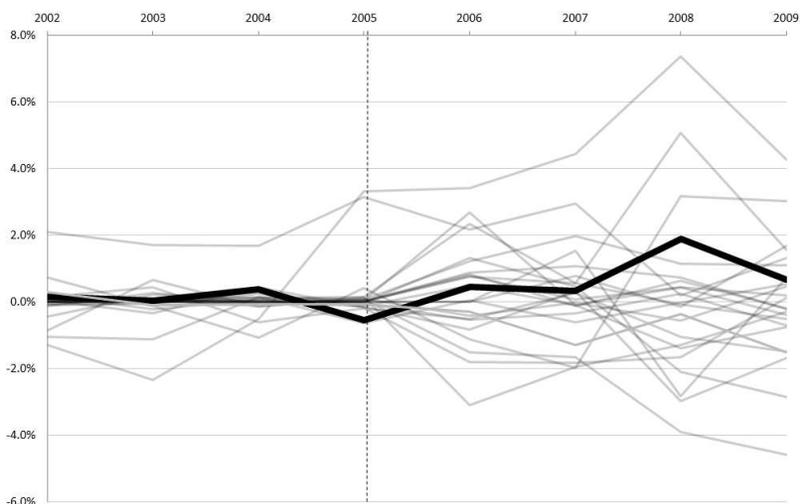
remuneração abaixo desses valores. Para uma visão mais completa do ajuste do mercado de trabalho, passa-se à análise dos demais indicadores mencionados na introdução, quais sejam, a taxa de emprego, a taxa de informalidade e o hiato salarial entre empregados com e sem carteira assinada.

## 5.2 Taxa de emprego

As Figuras 13 e 14 mostram os resultados das estimativas de efeito dos pisos salariais do PR e de SP sobre a taxa de emprego, definida como a razão entre o total de empregados nos grupos com piso salarial e a PIA na respectiva UF. Em ambos os casos, não há evidência de efeito. A Figura 13 apresenta as estimativas para o PR, que tendem a ser próximas de zero, com o procedimento de inferência reforçando este fato em quase todos os anos. A Figura 14 traz os resultados para SP, em que o ajuste do grupo de controle sintético no período pré-tratamento deixa a desejar, o que não permite tirar nenhuma conclusão sobre o efeito a ser estimado.

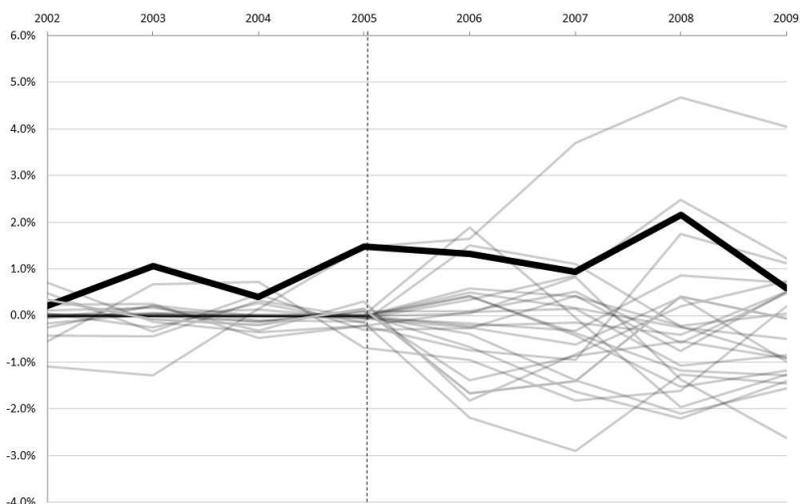
Tal como foi visto na subseção anterior, é possível que a agregação dos grupos ocupacionais dilua eventuais efeitos presentes em ao menos um grupo isoladamente. Sendo assim, as Figuras 15 e 16 reportam os efeitos sobre a taxa de emprego nos grupos ocupacionais 5 do PR e 1 de SP.<sup>28</sup> Em ambos os casos, há uma leve tendência de aumento desse indicador após a introdução dos pisos. As estimativas pontuais chegam a 0,7% da PIA do PR e a 0,5% da PIA de SP na média de 2007 a 2009. Em ambos os casos, são aumentos estimados de 6,5% no contingente empregado nessas ocupações em relação ao que se observaria na ausência dos pisos. Além de registrarem os maiores valores em cada um dos dois estados, os efeitos nesses grupos parecem ser os únicos com significância estatística.

<sup>28</sup>No caso das taxas de emprego calculadas para cada grupo ocupacional, vale notar que, enquanto o total de empregados é específico do grupo considerado, a PIA é agregada para a respectiva UF.



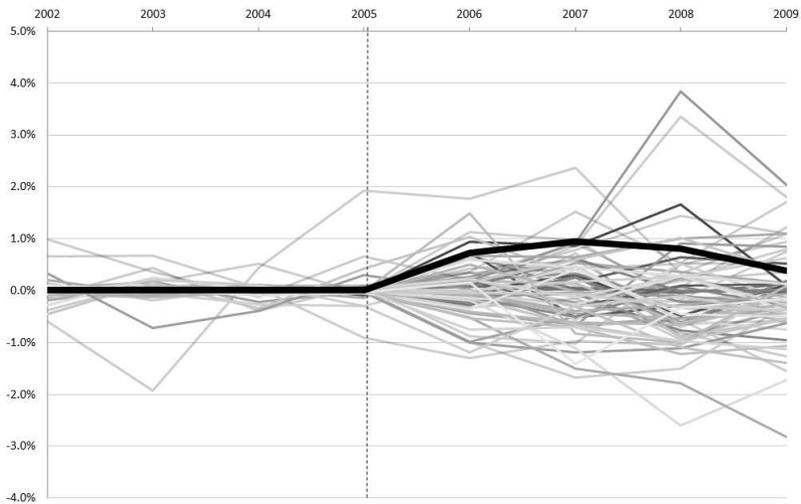
Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 13:** Empregados com piso/PIA: PR e demais UFs



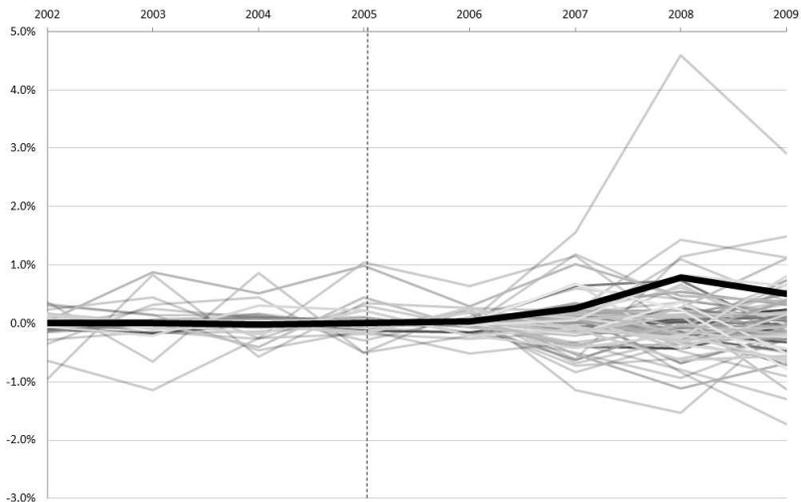
Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 14:** Empregados com piso/PIA: SP e demais UFs



Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 15:** Empregados/PIA: grupo 5 do PR e grupos das demais UFs



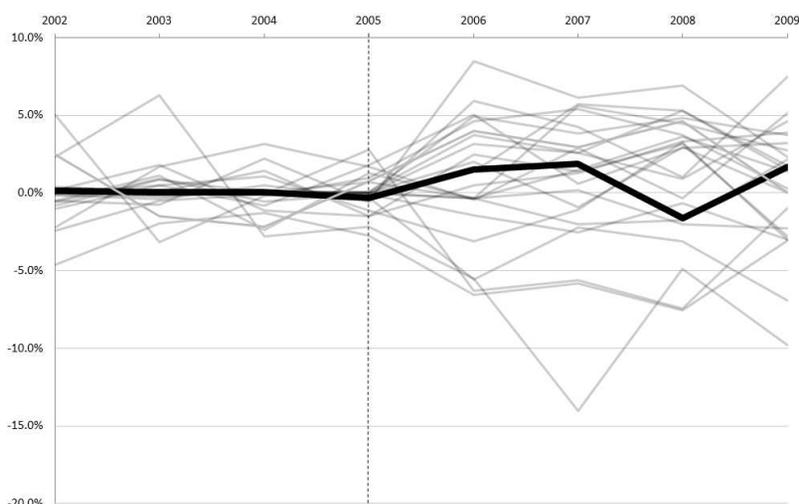
Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 16:** Empregados/PIA: grupo 1 de SP e grupos das demais UFs

Juntando-se as partes, as evidências são de que, ao menos em um grupo ocupacional de cada UF, a introdução do piso melhorou a distribuição de salários sem efeitos negativos sobre o emprego, que pode até ter reagido de forma positiva. Antes de avaliar a introdução dos pisos como benéfica para as ocupações em questão, resta saber se o aumento do emprego não foi fruto de um aumento no grau de informalidade.

### 5.3 Taxa de informalidade e hiato salarial dos com e sem carteira assinada

As Figuras 17 e 18 mostram os resultados para as taxas de informalidade entre empregados com e sem carteira de trabalho assinada no agregado das ocupações com piso no PR e em SP, respectivamente. No caso do PR, ilustrado na Figura 17, há uma clara indicação de que o piso salarial não teria afetado o grau de formalização dos empregados. A Figura 18 parece indicar uma tendência de declínio da informalidade para SP, ainda que não totalmente corroborada pelo método de inferência.



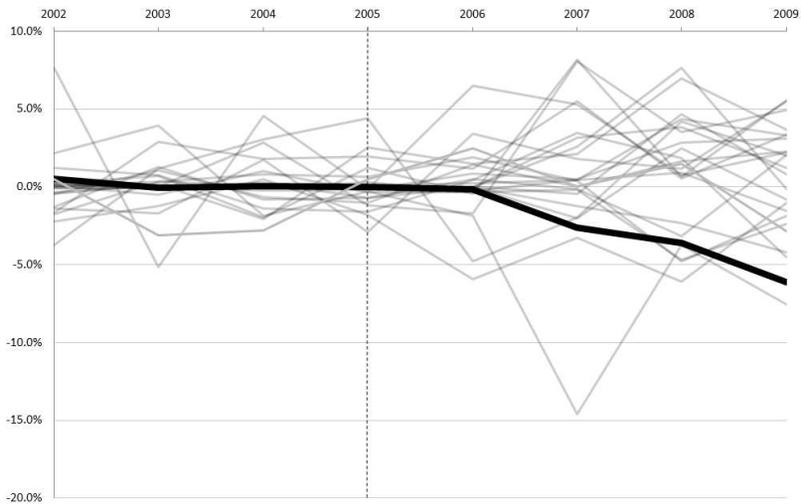
Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 17:** Informalidade: PR e demais UF's

As Figuras 19 e 20 reportam o efeito dos pisos sobre o hiato salarial entre empregados com e sem carteira no agregado das ocupações com piso salarial tanto no PR como em SP. Nesse caso, os resultados apontam inequivocamente para um efeito nulo.

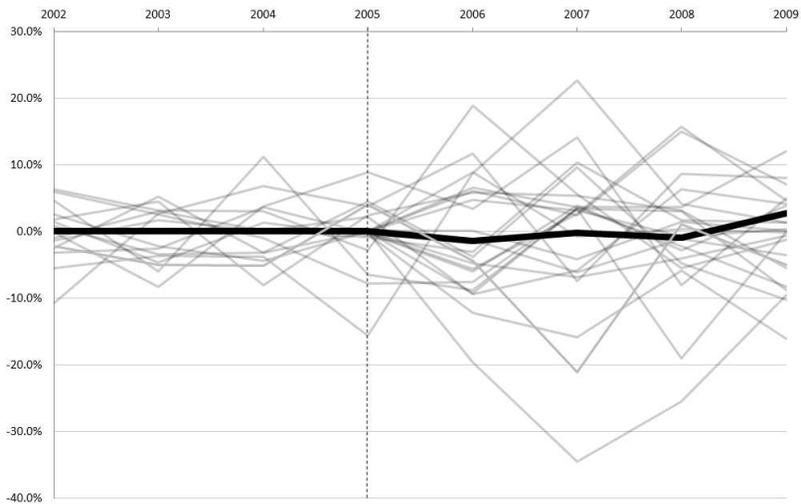
A análise dos efeitos dos pisos sobre o grau de informalidade e o hiato salarial foi repetida para os grupos ocupacionais 5 do PR e 1 de SP. Embora não apresentados, os resultados mostram que os pisos parecem não ter afetado nenhum dos dois indicadores dos grupos ocupacionais analisados.<sup>29</sup>

<sup>29</sup>Esses resultados podem ser requisitados aos autores.



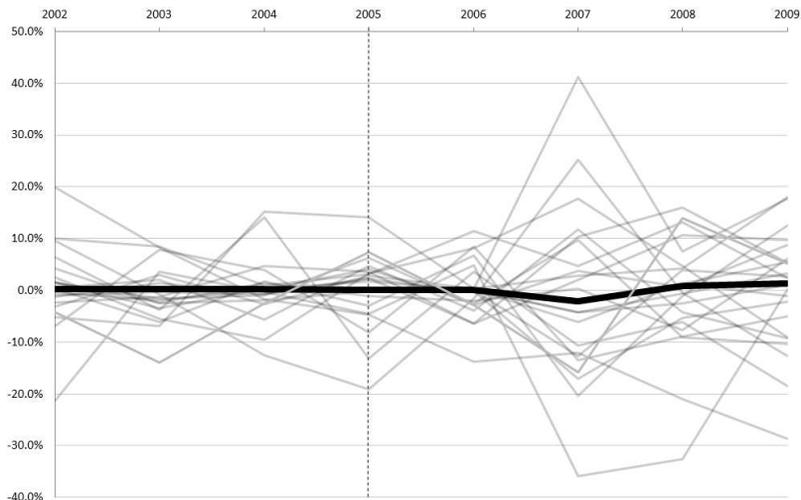
Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 18:** Informalidade: SP e demais UFs



Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 19:** Hiato salarial formal-informal: PR e demais UFs



Fonte: PNAD/IBGE Elaboração dos autores.

**Figura 20:** Hiato salarial formal-informal: SP e demais UFs

#### 5.4 Discussão dos resultados

Esta subseção discute, de forma especulativa, possíveis determinantes dos resultados descritos anteriormente. Estes resultados podem ser resumidos da seguinte forma. Tanto no PR como em SP, a introdução de pisos salariais teve efeitos heterogêneos entre os grupos ocupacionais analisados.<sup>30</sup> Em ambas as UFs, dois dos três grupos ocupacionais considerados não apresentaram efeitos significativos da introdução dos pisos sobre nenhum dos indicadores analisados. Esse fato seria condizente com uma baixa aderência dos contratos de trabalho aos valores legislados para esses grupos ocupacionais. Aparentemente, o trabalhador que seria remunerado em uma quantia inferior ao piso se ele não existisse permaneceu nesta situação na vigência do piso. O mesmo ocorreu com aquele que teria salário superior ao piso introduzido nessas respectivas ocupações. Não se descarta a hipótese de a lei ter sido ignorada para esses grupos ocupacionais.<sup>31</sup>

No entanto, os resultados do grupo ocupacional 5 do PR e do grupo ocupacional 1 de SP demonstram tanto algum grau de cumprimento da lei como um ajuste do mercado de trabalho dessas ocupações à introdução do piso. Os seguintes efeitos atribuídos à introdução do piso salarial foram documentados nestes grupos ocupacionais: *i*) diminuição da parcela de trabalhadores com remuneração abaixo do piso; *ii*) aumento da taxa de emprego, ao menos no segundo ano após a introdução do piso; *iii*) taxas de informalidade inalteradas; e *iv*) diferenciais de salários entre empregados com e sem carteira assinada inalterados.

Dados os efeitos apontados nos dois primeiros itens, a aparente falta de efeito nos itens (*iii*) e (*iv*) sugere que os ganhos advindos da introdução do piso para empregos e salários não se restringiram ao setor formal. Não só o

<sup>30</sup>Cabe lembrar que os grupos ocupacionais utilizados neste artigo para SP e PR encontram-se descritos no Apêndice A.

<sup>31</sup>Como mencionado na seção 2, Moura & Neri (2008) para o RJ e RS e Chahad & Pereda (2010) para SP também encontram baixa efetividade dos pisos salariais introduzidos nesses estados. O primeiro estudo também não encontra efeito do piso sobre o emprego.

setor informal parece ter se beneficiado nessas dimensões, como parece tê-lo feito em proporções similares ao setor formal. Essas especulações estão em linha com a tese do efeito-farol já documentado em trabalhos anteriores sobre SM no Brasil (Foguel 1998, Neri et al. 2001).

Dessa forma, parece que a introdução dos pisos para os dois grupos ocupacionais foi absorvida pelos empregadores dessas ocupações sem que estes reduzissem o emprego ou passassem a informalizar seus contratos de trabalho. Ou seja, nessas ocupações os legisladores aparentemente conseguiram o resultado desejado em uma parte dos postos, sem os efeitos colaterais mais temidos. Assim mesmo, vale lembrar que a parcela dos postos com remuneração abaixo do piso ainda é bem significativa, conforme exposto nas tabelas 2 e 3.

No que tange aos efeitos apontados sobre a melhora na distribuição de salários e o aumento na taxa de emprego, a literatura aponta três mecanismos alternativos compatíveis com tais resultados. O primeiro, mencionado na introdução, está relacionado aos modelos que atribuem algum poder de monopólio às firmas no mercado de trabalho. Essa classe de modelos inclui desde aqueles tradicionais em que o poder de monopólio é derivado de uma baixa densidade de firmas até os modelos de busca por emprego, em que o poder de monopólio decorre da imperfeição na informação que flui para os trabalhadores sobre postos vagos. Nessa classe de modelos, as firmas podem não alterar seu nível de emprego frente a um aumento do piso salarial vigente ou mesmo aumentá-lo caso haja um aumento na oferta de trabalhadores dispostos a trabalhar pelo valor do piso.

O segundo mecanismo relaciona-se ao arcabouço de salário-eficiência, tal como sugerido por Rebitzer & Taylor (1995). A principal característica deste arcabouço consiste em incorporar o nível de esforço que o trabalhador decide envidar na execução do seu ofício, em um ambiente em que a firma não consegue monitorá-lo de forma perfeita, mas o ameaça de demissão no caso de conseguir detectá-lo não se esforçando.<sup>32</sup> Os autores demonstram que, para um valor de piso mais alto que aquele que induz o pleno esforço dos trabalhadores, porém não demasiadamente alto, tanto a introdução de um piso quanto o seu aumento tendem a aumentar o emprego.<sup>33</sup>

O terceiro mecanismo advém de modelos que levam em consideração o caráter parcial do cumprimento do piso. Em alguns desses modelos, há predições que apontam para o aumento do emprego sob certas condições relativas à capacidade do Estado em fazer a lei ser respeitada (Yaniv 2004).

Por fim, efeitos positivos de pisos salariais sobre o emprego são justificados na literatura também por argumentos centrados no estímulo à demanda por produtos advindos do aumento de rendimentos auferidos por empregados afetados pelo piso.<sup>34</sup> Porém, no caso dos resultados aqui reportados, é preciso atentar que se observa um efeito positivo de emprego apenas para um grupo

<sup>32</sup>Por consequência, as firmas optam por pagar salários mais altos que aquele que equilibra o mercado de trabalho em pleno emprego. Dessa forma, os trabalhadores decidem sempre se esforçar.

<sup>33</sup>De forma intuitiva, o resultado se deve ao piso tornar desnecessário para a firma monitorar o esforço do trabalhador. Assim, os recursos empregados na atividade de monitoramento podem ser utilizados na ampliação da força de trabalho da firma.

<sup>34</sup>Um estudo interessante sobre impactos do SM que incorpora os efeitos de variação na renda dos trabalhadores é Fizbein (1992).

ocupacional específico em cada estado. Logo, não é óbvio que o argumento via demanda por produtos justifique os resultados encontrados.<sup>35</sup>

## 6 Conclusão

Este trabalho apresentou uma avaliação dos efeitos dos pisos estaduais introduzidos no PR e em SP. Em ambas as UFs, foram analisados três grupos ocupacionais distintos com base nos dados da PNAD do IBGE para o período de 2002 a 2009. Chega-se à conclusão de que, em um dos grupos ocupacionais de cada UF, o piso parece ter aumentado a remuneração de trabalhadores que recebiam abaixo dos respectivos valores, sem que houvesse efeitos de contração da taxa de emprego ou aumento da informalidade. Em particular, o emprego parece ter até reagido de forma positiva nessas ocupações. Foram mencionadas algumas interpretações alternativas desse resultado com base na literatura internacional, que também já se deparou com efeitos sobre o emprego na mesma direção. Ainda para os mesmos grupos ocupacionais, foram encontradas evidências referentes à informalidade que corroboram a tese do efeito-farol.

Nos demais grupos ocupacionais, as evidências apontam que a lei foi ignorada pelas partes nas relações empregatícias, pois continuaram sendo pagos salários inferiores aos pisos em uma proporção tão alta quanto a verificada onde eles não existem.

Uma das principais contribuições deste trabalho foi a utilização de uma nova metodologia para identificar efeitos de tratamento em unidades agregadas. Trata-se do método de grupo de controle sintético apresentado por Abadie et al. (2010), mais flexível e baseado em um modelo mais geral que o de diferenças em diferenças, comumente empregado na literatura. Acredita-se que esse método possa ser útil em diversas outras aplicações em que tenha havido um experimento natural decorrente de uma mudança inesperada em alguma política local.

## Referências Bibliográficas

- Abadie, A., Diamond, A. & Hainmueller, J. (2010), 'Synthetic control methods for comparative case studies: estimating the effect of california's tobacco control program', *Journal of the American Statistical Association* **105**(490).
- Abadie, A. & Gardeazabal, J. (2003), 'The economic costs of conflict: a case study of the basque country', *The American Economic Review* **93**(1).
- Angrist, J. D. & Krueger, A. B. (1999), Empirical strategies in labor economics, in O. Ashenfelter & D. E. Card, eds, 'Handbook of Labor Economics', 1. ed. Elsevier. v. 3.
- BLUNDELL, R. COSTA DIAS, M. (2009), 'Alternative approaches to evaluation in empirical microeconomics.', *Journal of Human Resources* **44**(3).

---

<sup>35</sup>Seria necessária alguma hipótese adicional. Por exemplo, de que a demanda pelos bens e serviços intensivos na mão de obra cujos salários foram elevados pelos pisos fosse especialmente elástica à renda dos trabalhadores empregados em sua própria produção.

- Brasil (1988), 'Constituição da república federativa do brasil de 1988'.
- Card, D. (1992), 'Do minimum wages reduce employment? a case study of california, 1987-1989', *Industrial and Labor Relations Review* **46**, 38–54.
- Card, D. & Krueger, A. B. (1994), 'Minimum wages and employment: a case study of the fast-food industry in new jersey and pennsylvania', *American Economic Review* **84**, 772–793.
- Card, D. & Krueger, A. B. (1995), *Myth and measurement: the new economics of the minimum wage*, Princeton: Princeton University Press.
- Chahad, J. P. Z. & Pereda, P. C. (2010), O comportamento e a avaliação da efetividade do piso salarial regional paulista, in 'ENCONTRO NACIONAL DE ECONOMIA, Bahia. Anais... Salvador: ANPEC'.
- Corseuil, C. H. & Servo, L. M. S. (2002), Salário mínimo e bem-estar social no brasil: uma resenha da literatura, Texto para Discussão 880, Ipea, Brasília.
- Cunningham, W. (2007), Minimum wages and social policy: lessons from developing countries, Technical report, Washington: World Bank.
- Danziger, L. (2009), 'Noncompliance and the effects of the minimum wage on hours and welfare in competitive labor markets', *Labour Economics* **16**(6), 625–630.
- Danziger, L. (2010), 'Endogenous monopsony and the perverse effect of the minimum wage in small firms', *Labour Economics* **17**, 224–229.
- Dube, A., Lester, W. & Reich, M. (2010), 'Minimum wage effects across state borders: estimates using contiguous counties', *The Review of Economics and Statistics* **92**(4), 945–964.
- Dube, A., Lester, W. & Reich, M. (2011), Do frictions matter in the labor market? accessions, separations and minimum wage effects, Discussion Paper 5.811, IZA, Bonn.
- Fajnzylber, P. (2001), Minimum wage effects throughout the wage distribution: evidence from brazil's formal and informal sectors, Texto para Discussão 151, Cedeplar, Belo Horizonte.
- Fizbein, A. (1992), o workers in the informal sector benefit from cuts in minimum wage?, Working Paper 826, World Bank, Washington.
- Foguel, M. (1998), Uma avaliação dos efeitos do salário mínimo sobre o mercado de trabalho no brasil, Texto para Discussão 564, Ipea. Rio de Janeiro.
- Garloff, A. A. (2010), 'Minimum wages, wage dispersion and unemployment in search models: a review', *Zeitschrift für Arbeitsmarkt Forschung* **43**, 145–167.
- Jatobá, J. & Chahad, J. P. Z. (1997), 'O papel do salário mínimo no contexto da estabilidade econômica', *Mercado de trabalho: conjuntura e análise* (n. 3, Ipea, Rio de Janeiro).
- Lemos, S. (2004), The effects of the minimum wage in the formal and informal sectors in brazil, Discussion Paper 1.089, IZA, Bonn.

- Masters, A. (1999), 'Wage posting in two-sided search and the minimum wage', *International Economic Review* **40**(4).
- Meyer, B. (1995), 'Natural and quasi-experiments in economics', *Journal of Business and Economic Statistics* **13**, 151–161.
- Mincer, J. (1976), 'Unemployment effects of minimum wages', *Journal of Political Economy* **84**, 87–105.
- Moura, R. L. & Neri, M. C. (2008), 'Efetividade do "salário mínimo estadual": uma análise via regressões quantílicas para dados longitudinais', *Economia Aplicada* **12**(2), 239–273.
- Neri, M. (1999), Salário mínimo: o reajuste de 99, a desvinculação e a regionalização, Technical Report 45, Boletim Conjuntural. Rio de Janeiro.
- Neri, M., Gonzaga, G. & Camargo, J. M. (2001), 'Salário mínimo, "efeito farol" e pobreza', *Revista de Economia Política* **21**, 78–90.
- Neumark, D. & Wascher, W. (2000), 'Minimum wages and employment: a case study of the fast-food industry in new jersey and pennsylvania: comment', *The American Economic Review* **90**, 1.362–1.396.
- Rebitzer, J. & Taylor, L. (1995), 'The consequences of minimum wage laws: some new theoretical ideas', *Journal of Public Economics* **56**(2), 245–255.
- Roppen, O. (2011), 'Reconciling the evidence of card and krueger (1994) and neumark and wascher (2000)', *Journal of Applied Econometrics* **26**, 1.051–1.057.
- Saboia, J. (1984), Evolução histórica do salário mínimo no brasil: fixação, valor real e diferenciação regional, PNPE - Série Fac-Símile 15, Ipea. Rio de Janeiro.
- Soares, F. V. (2004), Minimum wage and the informal sector in Brazil, Dissertation (philosophy doctor), University College London.
- Swinnerton, K. (1996), 'Minimum wages in an equilibrium search model with diminishing returns to labor in production', *Journal of Labor Economics* **14**(2), 340–355.
- Ulysea, G. & Foguel, M. N. (2006), Efeitos do salário mínimo sobre o mercado de trabalho brasileiro, Texto para Discussão 1.168, Ipea. Rio de Janeiro.
- Yaniv (2004), 'Minimum wage noncompliance and the sub-minimum wage rate', *Economics Bulletin* v. **10**(n. 9), p. 1–7.

## Apêndice A

**Tabela A.1:** Composição dos grupos ocupacionais analisados no PR

Grupo	Piso <sup>(1)</sup>	Grandes grupos da CBO 2002 citados na lei a partir de 2007
5	II	GG 5 – Trabalhadores dos serviços, vendedores do comércio em lojas e mercados
	I	GG 6 – Trabalhadores agropecuários, florestais, da caça e pesca
6	III	GG 9 – Trabalhadores de manutenção e reparação
	IV	GG 4 – Trabalhadores de serviços administrativos
	VI	GG 3 – Técnicos de nível médio
8	V	GG 7 – Trabalhadores da produção de bens e serviços industriais (sistemas “discretos”)
		GG 8 – Trabalhadores da produção de bens e serviços industriais (sistemas “contínuos”)

Elaboração dos autores.

Nota: <sup>(1)</sup> Seis faixas de pisos vigentes de 2007 a 2009. Em 2006, a lei definiu seis faixas de pisos para uma lista de ocupações específicas, mas, nas estatísticas deste estudo, foi atribuído o menor piso daquele ano aos grupos denominados 5, 6 e 8 indistintamente. Depois do período coberto na análise, as faixas II, III e IV foram fundidas em uma só.

**Tabela A.2:** Composição dos grupos ocupacionais analisados em SP

Grupo	Piso	Ocupações específicas conforme listadas na lei <sup>(1)</sup>
1	I	Trabalhadores domésticos, serventes, trabalhadores agropecuários e florestais, pescadores, contínuos, mensageiros e trabalhadores de serviços de limpeza e conservação, trabalhadores de serviços de manutenção de áreas verdes e de logradouros públicos, auxiliares de serviços gerais de escritório, empregados não especializados do comércio, da indústria e de serviços administrativos, cumins, barboys, lavadeiros, ascensoristas, motoboys, trabalhadores de movimentação e manipulação de mercadorias e materiais e trabalhadores não especializados de minas e pedreiras.
2	II	Operadores de máquinas e implementos agrícolas e florestais, de máquinas da construção civil, de mineração e de cortar e lavar madeira, classificadores de correspondência e carteiros, tintureiros, barbeiros, cabeleireiros, manicures e pedicuros, dedetizadores, vendedores, trabalhadores de costura e estofadores, pedreiros, trabalhadores de preparação de alimentos e bebidas, de fabricação e confecção de papel e papelão, trabalhadores em serviços de proteção e segurança pessoal e patrimonial, trabalhadores de serviços de turismo e hospedagem, garçons, cobradores de transportes coletivos, barmen, pintores, encanadores, soldados, chapeadores, montadores de estruturas metálicas, vidreiros e ceramistas, fiandeiros, tecelões, tingidores, trabalhadores de curtimento, joalheiros, ourives, operadores de máquinas de escritório, datilógrafos, digitadores, telefonistas, operadores de telefone e de telemarketing, atendentes e comissários de serviços de transporte de passageiros, trabalhadores de redes de energia e de telecomunicações, mestres e contramestres, marceneiros, trabalhadores em usinagem de metais, ajustadores mecânicos, montadores de máquinas, operadores de instalações de processamento químico e supervisores de produção e manutenção industrial.
3	III	Administradores agropecuários e florestais, trabalhadores de serviços de higiene e saúde, chefes de serviços de transportes e de comunicações, supervisores de compras e de vendas, agentes técnicos em vendas e representantes comerciais, operadores de estação de rádio e de estação de televisão, de equipamentos de sonorização e de projeção cinematográfica e técnicos em eletrônica.

Elaboração dos autores.

Nota: <sup>(1)</sup> Os secretários, excluídos da faixa de piso II a partir de 2009, foram desconsiderados nas análises deste estudo. Os técnicos em eletrônica, excluídos da faixa de piso III depois do período analisado, foram considerados.

**Tabela A.3:** Valores e pesos de cada variável usada na minimização da diferença pré-tratamento para estimar o efeito sobre a porcentagem abaixo do piso no PR-agregado

Variável	PR	PR-sintético	Peso na matriz V
Trabalhadores com remuneração abaixo do piso deflacionado em 2002 (%)	50,9	50,8	21,6
Trabalhadores com remuneração abaixo do piso deflacionado em 2003 (%)	50,9	50,8	24,0
Trabalhadores com remuneração abaixo do piso deflacionado em 2004 (%)	47,8	47,7	25,8
Trabalhadores com remuneração abaixo do piso deflacionado em 2005 (%)	46,3	46,2	28,6
Subtotal 1: resultado pré-tratamento (%)			100,0
Renda domiciliar per capita (R\$)	648,39	692,03	7,69E-27
Escolaridade média (anos de estudo)	7,67	7,84	3,23E-26
Trabalhadores empregados na indústria (%)	24,5	20,0	7,22E-27
Trabalhadores homens (%)	55,4	56,4	1,53E-26
Trabalhadores de cor branca (%)	72,8	44,0	7,08E-26
Trabalhadores com até 24 anos de idade (%)	30,8	30,9	1,45E-27
Subtotal 2: covariadas			1,35E-25
Total (%)			100,0

**Tabela A.4:** Peso de cada UF na composição do PR-sintético para a porcentagem abaixo do piso

UF	Peso w (%)
AM	34,1
DF	33,4
SC	21,2
AP	2,9
MT	1,1
RO	0,9
GO	0,9
MS	0,8
MG	0,6
ES	0,6
AC	0,4
PA	0,4
TO	0,4
RR	0,3
PE	0,3
BA	0,3
MA	0,2
CE	0,2
RN	0,2
PB	0,2
AL	0,2
SE	0,2
PI	0,1
Total	100,0

Fonte: PNAD/IBGE. Elaboração dos autores.

**Tabela A.5:** Frequência amostral por ano nas unidades tratadas e doadoras

	Grupo	2002	2003	2004	2005	2006	2007	2008	2009
Unidade tratada: PR	5	1795	1770	1862	1788	1949	1983	1964	2015
	6	1113	1162	1271	1211	1223	1222	1361	1422
	8	1373	1214	1491	1418	1461	1475	1561	1579
	Total	4281	4146	4624	4417	4633	4680	4886	5016
Média das UFs doadoras	5	878	881	941	991	1022	1013	1007	1046
	6	441	449	475	507	528	542	559	574
	8	564	554	590	609	632	622	676	670
	Total	1883	1884	2005	2107	2183	2177	2242	2289
Mediana das UFs doadoras	5	543	559	652	652	627	639	609	707
	6	296	315	310	334	358	359	388	409
	8	348	322	371	367	358	357	413	399
	Total	1187	1196	1333	1353	1343	1355	1410	1515
Unidade tratada: SP	1	2960	3055	3049	3177	3331	3173	3194	3198
	2	2755	2845	3017	3038	3174	3249	3301	2893
	3	512	522	512	540	521	560	573	602
	Total	6227	6422	6578	6755	7026	6982	7068	6693
Média das UFs doadoras	1	592	598	638	673	681	676	650	694
	2	488	488	502	545	555	563	563	564
	3	71	70	74	78	82	87	87	91
	Total	1151	1156	1214	1296	1318	1326	1299	1349
Mediana das UFs doadoras	1	340	382	385	400	358	380	365	474
	2	301	296	312	392	414	383	370	378
	3	43	35	39	40	51	50	50	45
	Total	684	713	736	832	823	813	785	897



## ENDIVIDAMENTO AGRÍCOLA: QUÃO COMPROMETIDOS SÃO OS PRODUTORES DO POLO PETROLINA-JUAZEIRO FRENTE A SUAS DÍVIDAS?

EMANOEL DE SOUZA BARROS \*  
LEONARDO FERRAZ XAVIER †  
DANIELLE TAVARES PESSOA ‡  
TIAGO FARIAS SOBEL §

### Resumo

O presente artigo busca identificar os determinantes da inadimplência entre os produtores de frutas do Polo Petrolina-Juazeiro. Para tanto, foi utilizado um modelo logit ordenado em função do produtor “escolher” se tornar inadimplente, renegociar sua dívida ou ser adimplente com os setores público e privado. Os resultados mostram que os produtores têm cinco vezes mais chances de se tornarem inadimplentes com o setor público do que com o setor privado. Há divergências quanto aos fatores que explicam a inadimplência frente ao setor público e frente ao setor privado, o que aprofunda a compreensão do tema em questão.

**Palavras-chave:** Agricultura; Inadimplência; Petrolina-Juazeiro; Logit; Probit.

### Abstract

This paper is aimed identifying variables for the default on debts of fruit producers of Petrolina/Juazeiro, regarding both public and private sectors. For that, we use a logit model where the producers may “choose” not paying their debts, renegotiate their debts or repay their debts. The probability of producers who decide not to pay their debts to the public sector is over 50%. There are differences in the variables that explain the non-payment to the public sector and the private sector.

**Keywords:** Agriculture; Default; Petrolina-Juazeiro; Logit; Probit.

**JEL classification:** C14, C21, Q13

**DOI:** <http://dx.doi.org/10.1590/1413-8050/ea128281>

---

\* UFPE(CAA). E-mail: embarros2002@yahoo.fr

† UFRPE. E-mail: leonardoferraz@gmail.com

‡ Governo de Pernambuco. E-mail: danielle.economia@hotmail.com

§ UFPB. E-mail: tiagosobel@yahoo.com.br

## 1 Introdução

O objetivo do crédito agrícola consiste em (i) disponibilizar recursos financeiros para custeio da produção e investimentos, (ii) além da comercialização dos produtos no período pós-colheita. Segundo autores como Guanziroli (2007) e Ortega & Só (2005), tais incentivos impulsionam o desenvolvimento da atividade, que, por sua vez, desempenha importante papel na sociedade, associado à provisão de alimentos à população, à geração de emprego e renda, à fixação do homem no campo etc.

Contudo, como destaca uma série de estudos — tais como os de Costa et al. (2008), Cardoso (2005), Faria et al. (2008), Zylbersztajn (2005) — a atividade agrícola implica elevados riscos a seus produtores, tendo em vista a possibilidade de ocorrência de fatores exógenos à ela intrínsecos — tais como intempéries da natureza, doenças e pragas, etc. — capazes de afetar consideravelmente não só o volume produzido, como a sua qualificação e, como consequência, a efetivação dos acordos vigentes entre os agentes.

Nesse contexto, dentre uma série de ações públicas que podem ser direcionadas ao produtor rural — tais como garantia de preços mínimos, seguro rural, pesquisa, extensão e assistência técnica, incentivos à redução de preços de insumos, etc. — é possível destacar as de incentivo ao crédito rural, uma vez que, como destaca Trombin (2007), estas dão condições aos produtores de dispor de meios adequados — tais como insumos, equipamentos, instalações e tecnologias específicas, como fertilizantes e defensivos, variedades de plantas adaptadas, equipamentos de irrigação adequados, etc. — que os viabilizem minimizar os riscos descritos.

Contudo, inerente à provisão do crédito, está a possibilidade de inadimplência em seu pagamento, especialmente, em se tratando da atividade agrícola, dadas as suas incertezas intrínsecas citadas e, por esta razão, ao fato de ela estar historicamente ligada a uma tendência de ação paternalista do Estado brasileiro. Segundo autores como Antão & Campanholo (2011), Villa Verde (2000) e Costa (2007), esta realidade tende a gerar uma cultura que afeta negativamente o comprometimento do produtor agrícola brasileiro no que tange ao pagamento de suas dívidas, especialmente, junto ao setor público.

Inseridos nesse contexto encontram-se os produtores do Polo Petrolina-Juazeiro, por este se tratar de um dos mais importantes centros produtores de frutas do país, cujo desenvolvimento se deu “de cima para baixo”, sob forte influência do governo central. Como destacam autores como Franca (2004) e Sobel (2011), o Polo apresenta um histórico de intervencionismo, onde o governo federal normalmente forneceu uma série de serviços facilitados ao produtor local, dentre os quais destacam-se: o crédito rural e acesso barato à terra, insumos, máquinas e equipamentos. Tal cenário criou uma cultura local paternalista que, em certa medida, perdura (sobretudo no caso dos produtores familiares), o que, acredita-se, tende a influenciar negativamente o comprometimento destes diante dos seus compromissos — destacadamente, com o setor público —, afetando indicadores de inadimplência locais.

Nesse escopo, a presente pesquisa busca exatamente analisar a inadimplência dos produtores agrícolas localizados no Polo Petrolina-Juazeiro. Nesse aspecto, o estudo recai sobre os seguintes conceitos de inadimplência: i) a inadimplência dos produtores frente às dívidas contraídas junto a seus fornecedores de insumos; ii) a inadimplência dos produtores frente às dívidas de regularização fundiária de suas propriedades rurais; e iii) a inadimplência dos

produtores frente ao crédito tomado junto a instituições financeiras. Portanto, no presente estudo, o tema da inadimplência é abordado de forma ampla, ultrapassando o conceito clássico e relativamente debatido acerca da inadimplência ao crédito fornecido por bancos para a realização de investimentos, comercialização ou à cobertura do custeio produtivo.

Como foco do trabalho, busca-se apontar o quanto o comprometimento dos produtores relativo às dívidas tomadas junto ao setor público é inferior (ou superior) ao comprometimento dos mesmos referentes às dívidas tomadas junto ao setor privado. Nesse aspecto, espera-se que a análise indique que os produtores do Polo tendam a ser mais comprometidos em pagar as dívidas tomadas com o setor privado do que em pagar as dívidas com o setor público. Referida expectativa deriva de uma série de fatores pormenorizados em estudos como os de Gatti et al. (1993), Dias (1991), Villa Verde (2000), Costa (2007), Antão & Campanholo (2011) e CODEVASF (2007, 2008, 2009 e 2010). De maneira geral, esses fatores podem ser divididos em duas principais vertentes: i) o histórico cultural de que as dívidas públicas<sup>1</sup> serão perdoadas no futuro, o que configura um problema de principal-agente; e ii) as dívidas contraídas junto ao setor privado<sup>2</sup> são caracterizadas, normalmente, como dívidas de curto prazo, o que leva o produtor a ser mais comprometido em efetuar seus pagamentos, sob o risco de impedir a continuidade de suas explorações produtivas.

Nesse sentido, as questões básicas em torno da análise do presente trabalho se encontram nos seguintes pontos:

- Qual a dimensão da inadimplência dos produtores no Polo Petrolina-Juazeiro?
- Quais fatores determinam referido fenômeno?
- Quanto os produtores do Polo são mais comprometidos em pagar as dívidas contraídas com o setor privado do que em pagar as dívidas contraídas junto a instituições públicas?
- Há diferenças quanto aos fatores que determinam a inadimplência junto ao setor privado com relação aos que determinam a inadimplência junto ao setor público?

Com vistas a um direcionamento objetivo aos resultados e conclusões do artigo, o mesmo foi organizado em cinco seções, incluindo a presente introdução. Na seção 2 tem-se a revisão da literatura ligada ao tema adimplência/inadimplência na ótica do mercado agrícola. Na Seção 3 “Metodologia”, apresentam-se os seguintes tópicos: i) modelo econométrico, com breve descrição do modelo teórico utilizado na pesquisa; ii) dados utilizados, que traz

---

<sup>1</sup>No escopo da presente pesquisa, as instituições tratadas como do setor público referem-se a: i) CODEVASF, instituição responsável pelo recolhimento das dívidas de regularização fundiária nos Perímetros Irrigados do Polo; e ii) bancos públicos, responsáveis pela concessão de financiamentos rurais.

<sup>2</sup>No que concerne aos produtores agrícolas do Polo Petrolina-Juazeiro, as dívidas contraídas junto ao setor privado recaem basicamente sobre o fornecimento de insumos. Dessa maneira, as principais instituições aqui contempladas como do setor privado referem-se a: i) casas comerciais, referentes ao fornecimento de insumos como sementes, adubos, defensivos etc.; e ii) Distrito de Irrigação, instituição privada sem fins lucrativos (associação civil) que fornece a água necessária para a irrigação das culturas exploradas.

informações acerca da amostra estudada; e iii) modelo empírico, em que se apresenta a estratégia de análise. Em seguida, os resultados são apresentados na Seção 4, a qual descreve características particulares aos modelos estimados, bem como traz diagnósticos de ajuste dos mesmos. As principais conclusões do trabalho são apresentadas na Seção 5.

## 2 Revisão da Literatura

É comum a literatura caracterizar uma relação contratual entre credor e devedor como uma relação com informação assimétrica — em que apenas uma das partes dispõe de informação imprescindível ao negócio —, sobretudo quando se trata do setor agrícola, dadas as incertezas inerentes da atividade. Nesse contexto, baseado em Chaddad & Lazzarini (2003), é possível identificar dois tipos comuns de problemas a partir das relações contratuais.

Sob uma ótica pré-contratual, é presumível que o devedor omite más informações ao credor com o objetivo de aumentar sua chance de obter o crédito a um custo menor. Sabendo dessa possibilidade, Stiglitz & Weiss (1981) asinam que o credor tende a elevar o prêmio exigido, o que acaba atraindo “devedores” de maior risco — pois sob esse cenário apenas projetos de maior retorno esperado mantêm-se atraentes —, gerando efeitos adversos sobre a inadimplência<sup>3</sup>. Já sob uma ótica pós-contratual, é também possível que o devedor, após o fechamento do contrato, tenha incentivos em alterar seu comportamento, realizando ações que também elevam o risco de inadimplência<sup>4</sup>.

Portanto, em razão de ambos os problemas potenciais, credores buscam realizar ações que os viabilizem obter o máximo de informações possíveis sobre o perfil do potencial devedor — por exemplo, analisando o seu histórico de créditos anteriores, seus dados financeiros e contábeis e, especialmente no mercado agrícola, seu perfil tecnológico — visando minimizar eventuais efeitos negativos da informação assimétrica, destacadamente, os riscos de inadimplência. Contudo, tais ações incorrem em custos, os quais são denominados de custos de informação (Stiglitz 2000). E, como apontam Azevedo & Shikida (2004), quanto maior a capacidade do tomador do crédito de esconder do credor o seu potencial de quitação de dívidas, maiores tendem a ser os custos de informação. Entretanto, embora tais ações minimizem o risco de inadimplência, não o elimina por completo, uma vez que agentes oportunistas têm incentivos em formar boa reputação no mercado, “se passando” por agentes honestos.

Assim, tornam-se necessários mecanismos mais complexos de auxílio no cumprimento dos contratos, redundando no que Williamson (1995) denominou de custos de transação. Esses custos podem ser de dois tipos: (i) pré-contratuais, tais como de negociação, de busca por salvaguardas e punições, de elaboração de contratos, com o tempo despendido para entendimento de legislações e cláusulas contratuais, entre outros; e (ii) pós-contratuais, tais como de monitoramento, fiscalização e controle de desempenhos, punições, renegociação de cláusulas contratuais, entre outros. Como apontam Chaddad & Lazzarini (2003), estes últimos decorrem das dificuldades de se prever todas as possíveis contingências futuras na fase pré-contratual — dada a existência de racionalidade limitada e oportunismo —, fazendo, portanto, com que

---

<sup>3</sup>Este caso é conhecido na literatura como seleção adversa (Akerlof 1970).

<sup>4</sup>Este caso é conhecido na literatura como risco moral (Arrow 1971).

os contratos sejam inerentemente incompletos e os mecanismos de controle de inadimplência inerentemente imperfeitos<sup>5</sup>.

Dado esse contexto, sabe-se que operações de crédito e a crédito junto aos produtores agrícolas no Brasil são realizadas por uma série de casas comerciais e instituições financeiras públicas e privadas, em que cada uma delas possui funções particulares no processo de fomento da atividade – seja concedendo financiamentos bancários, seja através de vendas de máquinas e insumos a prazo, entre outros. Nesse contexto, é interessante analisar algumas características gerais desses agentes, que influenciam no nível de inadimplência do produtor.

Primeiramente, sabe-se que bancos comerciais, públicos e privados, contam com depósitos de clientes como fonte de recursos, o que os viabilizam recursos estáveis e relativamente baratos para diferentes tipos de financiamentos. Contudo, segundo Chaddad & Lazzarini (2003), como normalmente não atuam diretamente em mercados agrícolas específicos, eles acabam tendo que gastar uma quantidade relativamente alta de tempo e recursos, na tentativa de reduzir a assimetria de informação junto aos produtores, visando minimizar o potencial de inadimplência. Porém, segundo Staten et al. (1990), produtores honestos e com adequada sanidade financeira podem auxiliar na minimização desses custos de informação, aceitando se submeter a processos rigorosos de avaliação de crédito sinalizando, dessa forma, sua “qualidade” e, consequentemente, podendo obter taxas de juros mais baixas — vale raciocínio similarmente inverso para produtores com alta probabilidade de inadimplência. Nesses termos, como apontam estudos como os de Antão & Campanholo (2011), Villa Verde (2000) e Costa (2007), estes últimos tendem a buscar credores menos rigorosos, particularmente no Brasil, entes públicos – através de bancos de desenvolvimento e de fomento regionais, linhas de créditos específicas (tais como o Pronaf-crédito<sup>6</sup>) e políticas públicas.

Por outro lado, agentes não financeiros que atuam nos Sistemas Agroindustriais — tais como empresas fornecedoras de máquinas, equipamentos e insumos — normalmente enfrentam maiores restrições relativas para viabilizar operações de crédito ao setor agrícola, por não dispor de uma fonte de recursos tão facilitada, devendo, portanto, buscar fontes alternativas – cujas principais, segundo Sherrick et al. (1994), acabam sendo os ativos líquidos da própria empresa, emissão de papéis comerciais e securitização de contas a receber. Por outro lado, por estes apresentarem relações comerciais mais íntimas e frequentes com os produtores, Chaddad & Lazzarini (2003) destacam que eles tendem a incorrer em menores custos de obter informações sobre o perfil tecnológico e a reputação do produtor, viabilizando, por essa via, um menor risco de inadimplência. Ademais, segundo Lazzarini et al. (1999), tais agentes podem reduzir ainda mais os riscos de inadimplência ao realizar contratos interligados com o produtor – por exemplo, por meio de venda de equipamentos atrelada à oferta de um serviço permanente de fornecimento de insu-

<sup>5</sup>Como aponta Rocha Júnior (2001), sempre existirá uma falha ou lacuna pela própria complexidade dos contratos e limitação do ser humano.

<sup>6</sup>Criado em 1996, o Programa Nacional de Fortalecimento de Agricultura Familiar (Pronaf) é um Programa de apoio ao desenvolvimento rural, a partir do fortalecimento da agricultura familiar. O PRONAF apresenta três linhas principais de atuação: PRONAF Crédito; PRONAF Infraestrutura e Serviço e o PRONAF Capacitação. O PRONAF Crédito é, dentre estas, a que mais libera recursos, financiando Custeio e Investimento relacionados às atividades agropecuárias, de pesca, de aquicultura, de extrativismo; e relacionados às atividades não-agrícolas, como o artesanato e o turismo rural (PRONAF 2006).

mos e/ou de transferência tecnológica. Esse tipo de contrato possibilita maior controle contra ações oportunistas, pois viabiliza (i) uma melhoria no desempenho técnico-econômico do produtor; e (ii) um maior comprometimento do devedor, visando a manutenção da parceria.

Por sua vez, segundo Chaddad & Lazzarini (2003), programas governamentais de crédito agrícola são os que possuem as maiores vantagens no que se refere ao custos de captação dos recursos para empréstimos, já que, no geral, tais programas utilizam recursos diretamente do tesouro nacional, que se financia por meio de tributos e de emissão de títulos de dívida pública, os quais apresentam menores riscos em relação aos emitidos por empresas privadas e, portanto, um custo relativamente mais baixo. Por outro lado, autores como Antão & Campanholo (2011), Villa Verde (2000) e Costa (2007) destacam que, nesse caso, há uma tendência a que os produtores agrícolas sejam menos comprometidos em pagar as dívidas, no geral, em razão do histórico cultural de que as dívidas públicas serão perdoadas no futuro, bem como pelo fato de as dívidas contraídas junto ao setor público serem, normalmente, caracterizadas como de longo prazo, o que leva o produtor a ser, relativamente, menos comprometido em efetuar seus pagamentos, sob o risco de impedir a continuidade de suas explorações produtivas.

### 3 Metodologia

O modelo de resposta binária é um caso especial de modelos com variáveis dependentes discretas, caso normalmente aplicado a estudos de inadimplência, cujo objetivo volta-se a apontar os determinantes que levam um produtor a tornar-se ou não inadimplente frente às dívidas tomadas. Nesse caso, assumem-se apenas dois valores para a resposta aos regressores do modelo: zero ou um. De acordo com Davidson & MacKinnon (2003), um modelo de resposta binária tenta explicar a probabilidade de o agente “escolher” a alternativa 1 como função de algumas variáveis explicativas observáveis.

Contudo, no presente estudo, além de averiguar a probabilidade de o agente “escolher” a inadimplência (alternativa 1), levar-se-á em consideração outro tipo de “escolha”: a opção de renegociar a dívida. Tal conduta estaria em uma situação intermediária entre a adimplência e a inadimplência, visto que: i) a renegociação revelaria uma dificuldade ou descompromisso do produtor em pagar sua dívida, o que é uma situação menos desejada que a adimplência; e ii) contudo, ao mesmo tempo, a renegociação demonstraria um interesse do produtor em quitar seu débito, o que se configura numa situação mais favorável que o efetivo descompromisso com a dívida, ou seja, a inadimplência de fato.

Nesse contexto, cuja variável dependente caracteriza-se como discreta e qualitativa, pretende-se analisar a probabilidade de o agente “escolher” uma dentre três alternativas (adimplência, renegociação da dívida ou inadimplência) como função de fatores observáveis. Para tal tipo de variável dependente, é adequada a aplicação do Logit Ordenado ou do Probit Ordenado, os quais utilizam máxima verossimilhança para fornecer previsões sobre quais fatores influenciam, no caso deste trabalho, a escolha da conduta do produtor frente a suas dívidas contraídas. Segundo Greene (2003), há vários exemplos de variáveis de escolha multinomial que estão inerentemente em ordem, como ratings de títulos, resultados de teste de gosto, pesquisas de opinião, nível de

cobertura de seguro tomado por um consumidor (nenhum, parcial ou total), emprego (desemprego, tempo parcial e tempo total), entre outros. Em cada um desses casos, embora o resultado seja discreto, o Logit Multinomial ou o Probit Multinomial podem falhar por não contabilizar a natureza ordinal da variável dependente. Assim, o modelo de regressão logística ordinal ou o Probit Ordenado apresentam-se como alternativas para referidas análises.

Teoricamente, a única real diferença entre os modelos Probit e Logit é a forma como os elementos dos parâmetros  $\beta$  são escalonados, devido às diferenças entre as variâncias das equações normal-padrão (Probit) e logística (Logit). Na prática, entretanto, ambos os modelos tendem a apresentar resultados bastante similares em relação às probabilidades estimadas.

### 3.1 Modelo Econométrico

Os modelos Logit Ordenado e Probit Ordenado são construídos a partir de uma regressão latente equivalente aos modelos binomiais tradicionais. A partir da função  $Y^* = X'\beta + \epsilon$  não há condições de observar  $Y$ , mas é possível observar que:

$$Y = \begin{cases} 0, \wedge Y^* \leq \mu_1 \\ 1, \wedge \mu_1 \leq Y^* \leq \mu_2 \\ 2, \wedge \mu_2 \leq Y^* \leq \mu_3 \\ \dots, J, \wedge \mu_J \leq Y^* \end{cases} \quad (1)$$

Em que os parâmetros  $\mu$  são chamados pontos de corte das categorias, sendo desconhecidos e estimados em conjunto ao vetor  $\beta$ , enquanto  $J$  refere-se ao número de categorias cuja variável dependente está dividida. As probabilidades de ocorrência em  $Y$  são então estimadas da seguinte maneira:

Em que a função  $F(\cdot)$  corresponde à função de densidade logística cumulativa, no caso dos modelos Logit; e à função de densidade normal-padrão cumulativa, no caso dos modelos Probit. As derivadas dessas probabilidades com relação a  $X$  correspondem, portanto, aos efeitos marginais de mudanças no regressor.

Para o ajuste do modelo, STATA (2012) sugere o teste razão de verossimilhança (teste LR) que verifica se pelo menos um dos coeficientes da regressão não é igual a zero no modelo. O teste pode ser calculado como  $-2[l(mod1) - l(mod2)]$ , em que  $l(mod1)$  é o logaritmo da probabilidade do modelo sem parâmetros e  $l(mod2)$  é o logaritmo da probabilidade do modelo com todos os parâmetros. O teste tem distribuição  $\chi^2$ , cujos graus de liberdade são definidos pelo número de regressores no modelo. A hipótese nula do modelo é que não existem previsores.

Outro tipo de teste de ajuste de modelo é o que analisa a presença de proporcionalidade entre os pontos de corte, ou seja, o teste da razão aproximada da verossimilhança (*likelihood ratio test*) proposto por Wolfe & Gould (1998). A hipótese nula é que não existe diferença nos coeficientes entre os modelos. O teste é semelhante ao teste LR, mas difere no número de graus de liberdade, que é igual a  $\rho(J-2)$ , em que  $\rho$  é o número de regressores e  $J$  é o número de categorias. A distribuição utilizada também é  $\chi^2$ . Esse teste também é sugerido para identificar a presença de heteroscedasticidade, como apontam Wang & Kockelman (2005).

Cabe ainda salientar o teste de Brant (1990), o qual analisa a hipótese de regressões paralelas. Referido teste é importante para a especificação do modelo, visto que, caso referida hipótese seja rejeitada, deve-se optar por um modelo logit ordenado generalizado. Se ainda assim a hipótese não for aceita, sugere-se como alternativa a estimação de um logit de chances proporcionais parciais.

### 3.2 Dados Utilizados

A amostra utilizada parte de uma pesquisa de campo realizada no primeiro trimestre de 2009 e gentilmente cedida pela Plantec Planejamento e Engenharia Agrônômica Ltda. A referida empresa é responsável pelos serviços de assistência técnica e extensão rural nos Perímetros Senador Nilo Coelho, Maria Tereza e Bebedouro<sup>7</sup>, conforme contrato estabelecido junto à Companhia de Desenvolvimento dos Vales do São Francisco e do Parnaíba (CODEVASF). A pesquisa contempla informações particulares a 1940 lotes localizados nos perímetros de irrigação acima citados.

As informações relativas à inadimplência dos produtores, tratadas na pesquisa de campo, contemplam quatro instituições: i) bancos públicos, relativo ao crédito tomado pelos produtores para a realização de investimentos, comercialização e/ou custeio produtivo; ii) CODEVASF, referente às dívidas de regularização fundiária das propriedades; iii) Distrito de Irrigação, relativo às dívidas de fornecimento de água para irrigação das explorações; e iv) casas comerciais, referente às dívidas de fornecimento de insumos contraídas junto aos fornecedores. Para os quatro casos, questões foram levantadas quanto à situação da dívida, ou seja, se o produtor encontra-se em situação de adimplência, renegociação da dívida ou inadimplência.

Vale destacar que a situação da dívida, neste trabalho, é tomada diante das seguintes hipóteses: i) um produtor com dívidas junto aos bancos públicos só pode ser assim tratado se o mesmo contraiu algum tipo de financiamento<sup>8</sup>; ii) todos os produtores da amostra estão ou estiveram com dívidas junto à CODEVASF — assim, um produtor que não está em dívida junto a essa instituição será tratado como adimplente, enquanto aqueles que ainda estão em dívida poderão ser tratados como adimplentes, renegociantes ou inadimplentes, a depender de sua resposta —; iii) todos os produtores têm dívidas junto ao Distrito de Irrigação, visto que todos produzem e, portanto, participam do rateio de fornecimento de água feito por essa instituição; iv) todos os produtores têm dívidas junto às casas comerciais e, se compram à vista, serão aqui tratados como adimplentes, visto que pagaram no ato da compra. Dessa maneira, todos os produtores da amostra<sup>9</sup> serão tratados em uma das situações aqui analisadas — adimplência, renegociação da dívida ou inadimplência —, de acordo com a resposta dada por estes em cada questionário.

---

<sup>7</sup>A área produtiva do Polo Petrolina-Juazeiro é dividida em perímetros irrigados geridos pela CODEVASF. Os Perímetros Senador Nilo Coelho, Maria Tereza e Bebedouro somam mais de 55% das áreas irrigadas de todos os perímetros do Polo, mais de 75% dos colonos que atuam nos mesmos e quase 40% das empresas produtoras (CODEVASF 2012).

<sup>8</sup>Assim, produtores que nunca contraíram financiamento junto aos bancos públicos serão excluídos da amostra quando essa instituição for analisada.

<sup>9</sup>Exceto no caso dos que nunca contraíram financiamento junto aos bancos públicos, pelas razões acima destacadas.

### 3.3 Modelo Empírico

Com o objetivo de estimar a probabilidade de que o produtor do Polo Petrolina-Juazeiro, dadas determinadas características, possa vir a ser um potencial tomador de dívidas adimplente, renegociante ou inadimplente, considera-se como estratégia empírica a utilização de um modelo com variável dependente qualitativa ordinal. Ou seja, tem-se como resposta uma variável que assume o valor 0 para o produtor adimplente, o valor 1 para o renegociante e o valor 2 caso o mesmo esteja na situação de inadimplência.

Para responder ao questionamento sobre os fatores que determinam o fenômeno da inadimplência no Polo, a variável dependente utilizada poderá assumir os seguintes valores:

- 0, caso o produtor esteja em situação de adimplência nas quatro instituições aqui tratadas (bancos públicos, CODEVASE, Distrito de Irrigação e casas comerciais);
- 1, caso o produtor esteja em situação de renegociação da dívida em pelo menos uma das quatro instituições aqui tratadas, desde que não esteja inadimplente em qualquer outra das quatro instituições;
- 2, caso o produtor esteja em situação de inadimplência em pelo menos uma das quatro instituições aqui tratadas.

Assim, a variável dependente assumirá valor 2 quando o produtor estiver inadimplente em uma das quatro instituições aqui tratadas, independente de sua situação da dívida com as outras três instituições. Se o produtor não estiver inadimplente em nenhuma das instituições, basta que esteja em processo de renegociação com ao menos uma para que a variável dependente assumo o valor 1. Por fim, a variável dependente assumirá valor 0 somente quando o produtor encontre-se em situação de adimplência com todas as instituições aqui tratadas.

Definida a estrutura da variável dependente, esta será regredida em função de um conjunto de fatores para se estabelecer quais características podem influenciar as chances de o produtor se tornar adimplente, renegociante ou inadimplente. Seguindo Costa (2007) e Antão & Campanholo (2011), acerca dos fatores capazes de explicar a inadimplência rural, o modelo empírico estimado deve incluir variáveis relevantes a fim de se determinar a situação das dívidas. A tabela 1 apresenta a lista de variáveis utilizadas para a estimação do modelo empírico básico, baseada nos trabalhos acima citados. Nesta, encontra-se a descrição das variáveis e seus valores conforme foi disposto nos questionários aplicados na pesquisa de campo.

As demais variáveis explicativas podem ser subdivididas em categorias que contemplam várias características do produtor ou do lote:

- Variáveis de localização: variáveis *dummy* para cada um dos onze Núcleos do Perímetro Nilo Coelho (N1 a N11)<sup>10</sup>, uma variável *dummy* para o Perímetro Maria Tereza e uma variável *dummy* para o Perímetro Bebedouro;

<sup>10</sup>O Perímetro Senador Nilo Coelho é dividido em onze núcleos distintos, com particularidades quanto às distâncias ao centro urbano de Petrolina e Juazeiro, perfil produtivo explorado, disponibilidade de mão de obra etc.

**Tabela 1:** Variáveis básicas utilizadas no modelo empírico

Variável	Descrição	Unidade / Valores Atribuídos
AREA	Área plantada	Hectares / Valores contínuos.
RENDA	Renda bruta mensal do lote	Faixa de renda / 0 (até R\$ 2 mil);1( R\$ 2 mil a R\$ 3 mil);2 (R\$ 3 mil a R\$ 4 mil); 3 (R\$ 4 mil a R\$ 5 mil);4 (R\$ 5 mil a R\$ 6 mil); 5 (R\$ 6 mil a R\$ 7 mil); 6 (além de R\$ 7 mil).
ESCOL	Escolaridade do produtor	Faixa de escolaridade / 0(analfabeto); 1 (alfabetizado);2 (Ensino Fundamental I); 3 (Ensino Fundamental II);4 (Ensino Médio ou Técnico); 5 (Ensino Superior).
IDADE	Idade do produtor	Faixa etária / 0 (até 30 anos); 1 (30 a 40 anos);2 (40 a 50 anos); 3 (50 a 60 anos); 4 (além de 60 anos).

Fonte:elaboração própria.

- Culturas exploradas na propriedade: sete variáveis correspondentes aos percentuais da área produzida cultivadas com manga, uva, banana, goiaba, acerola, coco e outras culturas;
- Tempo no lote: variável discreta categórica, referente ao tempo em que o produtor atua no Polo, variando de 0 a 4 segundo as faixas de 0 a 5 anos, de 6 a 10 anos, de 11 a 15 anos, de 16 a 20 anos e acima de 20 anos;
- Características do produtor: dez variáveis referentes a ter atividade em outro lote, número de lotes em que atua, ser colono, ser casado, ter filhos dependentes, ser do sexo feminino, ser portador de necessidades físicas, ter carro próprio, ter renda extra-rural e ter formação escolar na área agrícola;
- Características produtivas do lote: seis variáveis referentes a realizar análise de solo, realizar análise foliar, adotar recomendações de análises laboratoriais, adotar manejo específico para aumento do rendimento das culturas, contar com sistema de irrigação especificamente dimensionado para as culturas exploradas na propriedade e contar com certificação internacional (como Globalgap);
- Gestão do negócio: seis variáveis referentes ao número de funcionários, ter funcionário na administração do negócio, participar de treinamentos (produtor e/ou funcionários), participar de associação de produtores, realizar planejamento/controlar financeiro e realizar vendas sob contrato.

Dessa forma, a situação da dívida contraída pelo produtor diante das quatro instituições aqui tratadas será contemplada como variável dependente a ser regredida em função das diversas características aqui relacionadas. Assim, será possível responder ao questionamento inicialmente levantado sobre quais fatores determinam o fenômeno da inadimplência no Polo Petrolina-Juazeiro. Além disso, será possível estimar qual a probabilidade de um “produtor médio” do Polo<sup>11</sup> ser adimplente em suas dívidas, renegociante ou inadimplente de fato. Dessa forma, será possível apontar a dimensão da inadimplência no Polo, indicando quais as chances de seus produtores se tornarem adimplentes, renegociantes ou inadimplentes.

<sup>11</sup>O “produtor médio” caracteriza-se por valores das variáveis explicativas correspondentes às medianas amostrais.

Contudo, outros questionamentos relacionados aos objetivos deste trabalho ainda devem ser explorados, ou seja:

- Quanto os produtores do Polo são mais comprometidos em pagar as dívidas contraídas com o setor privado do que são em pagar as dívidas contraídas junto a instituições públicas?
- Há diferenças quanto aos fatores que determinam a inadimplência junto ao setor privado com relação aos que determinam a inadimplência junto ao setor público?

Para responder tais questionamentos, dois outros modelos serão estimados (doravante chamados modelos específicos), sendo a estrutura destes semelhante àquela apresentada até o momento (modelo conjunto, que leva em consideração as quatro instituições aqui tratadas). Nesse sentido, é necessário qualificar o que será tratado como dívidas contraídas com o setor público e como dívidas contraídas com o setor privado. Dentre as quatro instituições tratadas na pesquisa de campo, duas são qualificadas como dívidas com o setor público, quais sejam os bancos públicos e a CODEVASF; enquanto duas são relacionadas ao setor privado, quais sejam o Distrito de Irrigação e as casas comerciais.

Quanto ao modelo específico que contempla a análise do setor público, a estrutura da variável dependente tomará o mesmo formato do modelo conjunto, mas em vez de considerar as quatro instituições, levará em conta a situação da dívida apenas com os bancos públicos e com a CODEVASF. Dessa forma, a variável dependente assumirá valor 2 quando o produtor estiver inadimplente em uma dessas duas instituições, independente de sua situação da dívida com a outra; valor 1 quando o produtor estiver em renegociação em uma dessas instituições, desde que não esteja inadimplente na outra; e valor 0 quando o produtor estiver adimplente tanto com os bancos públicos, quanto com a CODEVASF. As variáveis explicativas, por sua vez, serão as mesmas relacionadas anteriormente, com respeito ao modelo conjunto.

De forma análoga, para o modelo específico que contempla a análise do setor privado, a estrutura da variável dependente levará em conta a situação da dívida apenas com o Distrito de Irrigação e as casas comerciais. Dessa forma, a variável dependente assumirá valor 2 quando o produtor estiver inadimplente em uma dessas duas instituições, independente de sua situação da dívida com a outra; valor 1 quando o produtor estiver em renegociação em uma dessas instituições, desde que não esteja inadimplente na outra; e valor 0 quando o produtor estiver adimplente tanto com o Distrito de Irrigação, quanto com as casas comerciais. Novamente, as variáveis explicativas serão as mesmas relacionadas anteriormente.

Dessa maneira, será possível identificar se existem diferenças com relação aos fatores que explicam a inadimplência dos produtores frente ao setor público e aqueles que explicam a inadimplência frente ao setor privado. Além disso, visto que é possível estimar a probabilidade de inadimplência do “produtor médio” do Polo em cada caso, pode-se dar apontamentos sobre o quanto os produtores são mais comprometidos em pagar dívidas com o setor privado do que são em pagar dívidas com o setor público.

## 4 Resultados

De acordo com Caouette et al. (1999), o risco de crédito está relacionado à capacidade de pagamento e características relacionadas ao caráter do devedor, sendo o primeiro fator o principal responsável pela inadimplência. Assim, conforme sugerem os autores, o modelo estimado deve considerar apenas variáveis com potencial de exercer influência sobre a capacidade de pagamento dos produtores rurais, considerando que incapacidade de pagamento está diretamente relacionada com a inadimplência, minimizando, portanto, a perspectiva do não pagamento ocorrer em função da vontade do devedor, aliando-se a isso a falta de elementos que viabilizem a análise sob esse ponto de vista.

Ademais, estimou-se os modelos Logit Ordenado e Probit Ordenado a fim de se verificar o efeito das variáveis expostas na Tabela 1 sobre a probabilidade de inadimplência do produtor rural no Polo Petrolina-Juazeiro. Para tanto, utilizou-se o método de máxima verossimilhança (ML), conforme apresentado no item referente à metodologia. Como referidos modelos são essencialmente heteroscedásticos, tornando os desvios-padrão estimados viesados, utilizou-se para a estimação da matriz de variância-covariância dos resíduos, desvios-padrão robustos à heteroscedasticidade. De acordo com Greene (2003), o uso de desvio-padrão robusto não altera os coeficientes estimados, entretanto, devido ao fato de que os erros-padrão são alterados, as estatísticas de teste fornecerão p-valores extremamente precisos.

### 4.1 Modelo Conjunto: Diagnóstico da Inadimplência Geral no Polo

Os resultados obtidos para o modelo básico são descritos na tabela 2 abaixo. As estatísticas LR  $\chi^2(4)$  e log pseudo-verossimilhança apontam para modelos apropriados, em que se verifica, tanto para o modelo Logit Ordenado, quanto para o Probit Ordenado, que a probabilidade de inadimplência tende a se elevar quanto maior o porte do produtor (captado pela variável AREA), enquanto tende a se reduzir quanto maior a renda obtida no lote, bem como a escolaridade e a idade do produtor. O número de observações utilizadas corresponde a 744 lotes, tendo em vista a ocorrência de *missing values* relativos às respostas sobre a situação da dívida dos produtores em cada uma das quatro instituições aqui tratadas (bancos públicos, CODEVASE, Distrito de Irrigação e casas comerciais), que formam a variável dependente; bem como *missing values* relativos às variáveis explicativas. Vale destacar, ainda, que a amostra utilizada na estimação desse modelo exclui os produtores que nunca contraíram crédito junto a bancos públicos, visto que esta é uma das instituições que dimensionam a variável dependente.

Ao se considerar a inclusão de outras variáveis ao modelo básico, conforme destacado na descrição metodológica do modelo empírico, verificou-se que variáveis auxiliares podem dar maior poder explicativo à probabilidade de inadimplência. Ao se testar o conjunto de variáveis auxiliares, o modelo de melhor ajuste aos dados amostrais é apresentado na tabela 3.

Novamente, as estatísticas LR  $\chi^2(10)$  e log pseudo-verossimilhança apontam para modelos apropriados. Tanto para o modelo Logit Ordenado, quanto para o Probit Ordenado, todas as variáveis mostraram-se significativas para explicar a situação das dívidas dos produtores da amostra - exceto no caso da *dummy* COLONO no modelo Logit -. Cabe destacar que, relativo ao modelo básico, apenas as variáveis RENDA e IDADE apresentaram instabilidade nos

**Tabela 2:** Estatísticas do modelo conjunto básico

Discriminação	Logit Ordenado		Probit Ordenado	
	Coefficientes	Ef. Marginais <sup>1</sup>	Coefficientes	Ef. Marginais <sup>1</sup>
AREA	0,07**	0,019**	0,04**	0,018**
RENDA	-0,24*	-0,061*	-0,14*	-0,057*
ESCOL	-0,20*	-0,050*	-0,13*	-0,050*
IDADE	-0,18*	-0,045*	-0,11*	-0,043*
Estatística LR $\chi^2(10)$	76,49*		79,39*	
Log pseudo-verossimilhança	-712,51		-712,59	
Nº de observações	744			

\* Significativo a 1%. \*\* Significativo a 5%. \*\*\* Significativo a 10%.

Variável dependente: categoria 0 — adimplente; categoria 1 — renegociante; categoria 2 — inadimplente.

<sup>1</sup> Trata-se dos efeitos marginais de alterações em cada variável explanatória sobre as chances de ocorrência na categoria 2 (aqui apresentados apenas para verificação da estabilidade dos modelos). Fonte: elaboração própria.

efeitos marginais, o que pode ser justificado pela inclusão de variáveis correlacionadas nos modelos de melhor ajuste. O número de observações utilizadas corresponde a 659 lotes, tendo em vista a ocorrência de *missing values*.

**Tabela 3:** Estatísticas do modelo conjunto de melhor ajuste

Discriminação	Logit Ordenado		Probit Ordenado	
	Coefficientes	Ef. Marginais <sup>1</sup>	Coefficientes	Ef. Marginais <sup>1</sup>
AREA	0,09*	0,022*	0,05**	0,020**
RENDA	-0,13*	-0,032*	-0,07*	-0,029*
ESCOL	-0,21*	-0,053*	-0,13*	-0,051*
IDADE	-0,33*	-0,083*	-0,20*	-0,079*
FILHOS	0,44***	0,106***	0,25***	0,097***
COLONO	0,29	0,073	0,21***	0,082***
NUMLOTES	0,08***	0,020***	0,05***	0,021***
CERTIF	-0,73*	-0,170*	-0,42*	-0,158*
PLANEJ	-0,18*	-0,044*	-0,11*	-0,045*
NUMFUNC	-0,05**	-0,012**	-0,03*	-0,011*
Estatística LR $\chi^2(10)$	103,63*		111,52*	
Log pseudo-verossimilhança	-609,61		-608,85	
Nº de observações	659			

\* Significativo a 1%. \*\* Significativo a 5%. \*\*\* Significativo a 10%.

Variável dependente: categoria 0 — adimplente; categoria 1 — renegociante; categoria 2 — inadimplente.

<sup>1</sup> Trata-se dos efeitos marginais de alterações em cada variável explanatória sobre as chances de ocorrência na categoria 2 (aqui apresentados apenas para verificação da estabilidade dos modelos). Fonte: elaboração própria.

Antes de analisar os resultados desse modelo, é interessante comentar as variáveis testadas que foram redundantes à identificação do mesmo<sup>12</sup>. Quanto às variáveis de localização, verificou-se que o simples fato de o produtor estar localizado em um ou outro Núcleo do Perímetro Nilo Coelho, ou nos Perímetros Maria Tereza e Bebedouro, não é capaz de explicar (de forma estatisticamente significativa) o comportamento dos produtores frente às dívidas contraídas. De forma análoga, o tipo de cultura ou culturas exploradas na propriedade também não explica referido comportamento.

<sup>12</sup>Para mais detalhes relacionados às variáveis auxiliares, vide descrição metodológica do modelo empírico.

O tempo no lote, que poderia captar a experiência do produtor com a atividade agrícola, também não chegou a explicar a situação das dívidas do produtor, assim como características como ser casado, ser do sexo feminino, ser portador de necessidades físicas, ter carro próprio, ter renda extra-rural e ter formação escolar na área agrícola. Contar com sistema de irrigação dimensionado especificamente para as culturas exploradas na propriedade também não se mostrou estatisticamente significativa. Outras variáveis, estas relacionadas à gestão do negócio, também não seriam capazes de explicar a variável dependente, quais sejam: ter funcionário na administração do negócio, participar de treinamentos, participar de associação de produtores e realizar vendas sob contrato.

De certa forma, esses resultados contrariam a hipótese de que a situação de inadimplência esteja relacionada com a capacidade de pagamento do devedor, visto que as variáveis descartadas estão aparentemente correlacionadas a tal capacidade. Contudo, os testes de redundância realizados sobre essas variáveis indicam que os fatores relacionados na tabela 3 são suficientes para explicar a situação das dívidas dos produtores da amostra. Para a análise desses fatores, é necessário, porém, apresentar os efeitos marginais referentes a cada variável.

Os efeitos marginais indicam o quanto à probabilidade de ocorrência em uma das categorias da variável dependente (adimplente, renegociante ou inadimplente) pode variar diante de alterações em determinada variável explanatória. Para tanto, é necessário estipular uma condição inicial referente às características do produtor, o que normalmente é feito sobre um “produtor médio”, que assumiria valores para as variáveis explicativas do modelo correspondentes à média ou à mediana amostral em cada característica<sup>13</sup>. Assim, a análise dos efeitos marginais parte de um choque em certa característica do “produtor médio” para se verificar as alterações na distribuição de probabilidades de ocorrência em uma das categorias da variável dependente.

Na amostra completa, de 1940 lotes, as medianas amostrais das variáveis explanatórias dispostas na tabela 3 correspondem a: AREA = 6,3 (área plantada no lote de 6,3 hectares); RENDA = 1 (renda bruta mensal do lote entre R\$ 2 mil e R\$ 3 mil); ESCOL = 4 (com Ensino Médio ou Técnico); IDADE = 3 (entre 50 e 60 anos); FILHOS = 1 (tem filhos dependentes); COLONO = 0 (é caracterizado como colono); NUMLOTES = 1 (atua em apenas um lote); CERTIF = 0 (não certificado); PLANEJ = 3 (faz planejamento de negócio de forma parcial-forte<sup>14</sup>); e NUMFUNC = 1 (conta com um único funcionário). Estas seriam, portanto, as características do “produtor médio” do Polo Petrolina Juazeiro.

Antes de explicar sobre os efeitos marginais de cada variável explicativa, é importante destacar que, de acordo com as estimativas realizadas a partir dos resultados da tabela 3, o “produtor médio” teria a seguinte distribuição de probabilidades de ocorrência em cada categoria da variável dependente<sup>15</sup>: chan-

<sup>13</sup>A opção entre média e mediana depende da distribuição da variável (na existência de *outliers*, p.ex., é preferível usar a mediana em vez da média), bem como do tipo de variável (para variáveis discretas, é preferível usar a mediana com vistas a manter a característica de variável discreta).

<sup>14</sup>A variável PLANEJ leva em consideração respostas a duas questões referentes às práticas de controle de custos e receitas, bem como pesquisa de preços de insumos. As respostas possíveis seriam “não realiza”, “realiza parcialmente” ou “realiza”. Assim, da combinação desses resultados, a variável PLANEJ foi criada variando de zero a quatro, de forma discreta e categórica, elevando-se conforme o grau de planejamento financeiro na gestão do negócio.

<sup>15</sup>Os resultados encontrados para essa distribuição mostraram que as estimativas dos modelos

ces de 14,64% de ser adimplente nas quatro instituições aqui tratadas (bancos comerciais, CODEVASF, Distrito de Irrigação e casas comerciais); 39,26% de renegociar dívidas com pelo menos uma das quatro instituições; e 46,09% de se tornar inadimplente com pelo menos uma das quatro instituições.

Esse resultado, portanto, é tratado como a dimensão do fenômeno da inadimplência no Polo Petrolina-Juazeiro, a qual representa níveis alarmantes quanto ao risco de crédito na região. Tais resultados ainda estão de acordo com as observações amostrais de que 18,1% dos produtores encontram-se na categoria de adimplência, 30,7% encontram-se na categoria de renegociante e 51,2% encontra-se no estrato de inadimplência.

Passando finalmente à análise dos fatores determinantes da inadimplência na região, de acordo com os resultados do modelo conjunto de melhor ajuste, vale ressaltar os seguintes pontos:

- Um choque de +1,0 ha sobre a área produzida (AREA = 7,3) leva o “produtor médio” a reduzir as chances de adimplência para 13,51%, reduzir as chances de renegociação para 38,38% e elevar as chances de inadimplência para 48,11%;
- Por sua vez, o aumento de uma faixa de renda (RENDA = 2) leva o mesmo a apresentar chances de 16,38%, 40,41% e 43,21% (respectivamente para adimplência, renegociação e inadimplência);
- O aumento de uma faixa de escolaridade (ESCOL = 5) altera a distribuição de probabilidades para 17,81% (adimplência), 41,18% (renegociação) e 41,01% (inadimplência);
- O aumento de uma faixa de idade (IDADE = 4) leva o “produtor médio” a variar sua distribuição para 19,69% (adimplência), 43,35% (renegociação) e 38,31% (inadimplência);
- Caso o mesmo não tivesse filhos dependentes (FILHOS = 0), sua distribuição passaria a ser de 21,15% (adimplência), 42,49% (renegociação) e 36,36% (inadimplência);
- Por sua vez, se o mesmo fosse caracterizado como colono (COLONO = 1), sua distribuição passaria para 10,41% (adimplência), 35,26% (renegociação) e 54,33% (inadimplência);
- Caso o “produtor médio” atuasse em mais um lote (LOTE = 2), sua distribuição passaria para 13,49% (adimplência), 38,36% (renegociação) e 48,15% (inadimplência);
- Se o mesmo fosse certificado (CERTIF = 1), sua distribuição de probabilidades passaria a ser de 26,34% (adimplência), 43,40% (renegociação) e 30,25% (inadimplência);
- Caso realizasse completamente práticas de gestão financeira (PLANEJ = 5), a distribuição seria de 17,40% (adimplência), 40,97% (renegociação) e 41,63% (inadimplência);

---

Logit e Probit Ordenado, conforme esperado, são praticamente idênticas. Contudo, visto que a estatística LR  $\chi^2(10)$  no modelo Probit é um pouco maior frente à mesma estatística no Logit, a análise dos efeitos marginais será realizada por meio das estimativas do Probit.

- Por fim, com a admissão de mais um funcionário no lote do “produtor médio” (NUMFUNC = 2), as chances seriam de 15,32% (adimplência), 39,74% (renegociação) e 44,95% (inadimplência).

A brevidade quanto aos comentários relativos a esses resultados deve-se à intenção de pormenorizar a análise relativa aos modelos específicos referentes aos setores público e privado. Ademais, demonstra-se que a inadimplência geral no Polo, analisada pela estimativa do modelo conjunto — relativo à situação da dívida dos produtores frente às instituições bancos públicos, CODEVASF, Distrito de Irrigação e casas comerciais —, relaciona-se positivamente com fatores como tamanho da propriedade, existência de filhos dependentes, ser caracterizado como colono e número de lotes em que o produtor atua. Por outro lado, as chances de adimplência tendem a se relacionar positivamente com fatores como renda do lote, escolaridade e idade do produtor, certificação produtiva (Globalgap), gestão financeira do negócio e número de funcionários no lote.

Entretanto, como descrito anteriormente, um ponto importante acerca dos fatores que influenciam a inadimplência no Polo pode estar relacionado ao tipo de instituição cujo produtor contrai suas dívidas. Nesse sentido, dois outros modelos são apresentados com vistas a qualificar o tipo de dívida, se contraída junto ao setor público ou se contraída junto ao setor privado. Nesse aspecto, espera-se que a distribuição de probabilidades nas categorias de situação das dívidas, bem como os fatores determinantes dessa ocorrência sejam divergentes segundo o tipo de dívida.

#### **4.2 Modelos Específicos: Diagnóstico da Inadimplência sobre Setores Público e Privado**

A estimação dos modelos específicos básicos encontra-se na tabela 4 a seguir<sup>16</sup>. Verifica-se que as estatísticas LR  $\chi^2(4)$  e log pseudo-verossimilhança apontam para modelos apropriados. Para o modelo específico relativo às dívidas com o setor público, os sinais dos coeficientes são equivalentes ao modelo conjunto básico apresentado na tabela 2, sendo semelhantes também os p-valores e efeitos marginais. Contudo, para o modelo específico relativo às dívidas com o setor privado, o modelo básico indica que as variáveis AREA, ESCOL e IDADE seriam estatisticamente insignificantes para explicar a inadimplência, o que evidencia a necessidade de se testar variáveis auxiliares.

Vale destacar que o número de observações varia em cada tipo de modelo. Isso ocorre porque, no caso do modelo específico relativo à dívida do setor público, sua variável dependente leva em consideração a situação da dívida frente às instituições bancos comerciais e CODEVASF; enquanto no caso do modelo para o setor privado, sua variável dependente leva em conta a situação da dívida com as instituições Distrito de Irrigação e casas comerciais. Além disso, cabe lembrar que, ao se considerar a situação da dívida com bancos comerciais para a formulação da variável dependente no modelo de setor

---

<sup>16</sup>Pelos mesmos motivos apresentados quanto à análise dos efeitos marginais do modelo conjunto, a apresentação dos resultados referentes aos modelos específicos (setor público e setor privado) será referente às estimativas do modelo Probit Ordenado.

**Tabela 4:** Estatísticas dos modelos específicos básicos

Discriminação	Dívida com Setor Público		Dívida com Setor Privado	
	Coefficientes	Ef. Marginais <sup>1</sup>	Coefficientes	Ef. Marginais <sup>1</sup>
AREA	0,04**	0,017**	0,02	0,004
RENDA	-0,15*	-0,059*	-0,07*	-0,013*
ESCOL	-0,11*	-0,043*	0,00	0,000
IDADE	-0,08**	-0,032**	-0,04	-0,008
Estatística LR	79,67*	11,36**		
Log pseudo-verossimilhança		-760,83		-799,36
Nº de observações		782		1030

\* Significativo a 1%. \*\* Significativo a 5%. \*\*\* Significativo a 10%.

Variável dependente: categoria 0 — adimplente; categoria 1 — renegociante; categoria 2 — inadimplente.

<sup>1</sup> Trata-se dos efeitos marginais de alterações em cada variável explanatória sobre as chances de ocorrência na categoria 2 (aqui apresentados apenas para verificação da estabilidade dos modelos). Fonte: elaboração própria.

público, o número de observações neste será ainda reduzido devido à exclusão daqueles entrevistados que nunca contraíram crédito para financiamento<sup>17</sup>.

Ao se considerar a inclusão de outras variáveis aos modelos específicos básicos, semelhante ao realizado na análise do modelo conjunto, verificou-se que variáveis auxiliares podem dar maior poder explicativo à probabilidade de inadimplência tanto relativa à dívida com o setor público, quanto com respeito à dívida com o setor privado.

Ao se testar o conjunto de variáveis auxiliares, o modelo de melhor ajuste aos dados amostrais utilizados para a estimação do modelo relativo ao setor público considera as variáveis AREA, RENDA, ESCOL, IDADE, COLONO, NUMLOTES, CERTIF, PLANEJ, NUMFUNC e CONDFUNC. Portanto, o modelo específico que vem a elucidar os fatores que influenciam a situação do produtor quanto a suas dívidas com o setor público é semelhante ao modelo conjunto, exceto pela ocorrência de impactos significativos causados pela variável CONDFUNC (*dummy* de lote conduzido por funcionário) e pela verificação de que a variável FILHOS (produtor com filho ou filhos dependentes) é estatisticamente insignificante para explicar a situação em que o produtor se encontra com respeito à dívida com o setor público. Além disso, os efeitos marginais de certas variáveis têm seus valores alterados consideravelmente, como será apresentado posteriormente com mais detalhe.

Por sua vez, o modelo de melhor ajuste relativo às dívidas com o setor privado demonstrou que a estrutura de fatores de influência é de fato divergente. Nesse aspecto, variáveis como AREA, ESCOL, COLONO, NUMLOTES, NUMFUNC e CONDFUNC são apontadas como redundantes ao modelo de melhor ajuste que, por sua vez, passa a considerar fatores que antes eram tratados como não explicativos, quais sejam RENDEXTR (*dummy* de renda extrarrural, seja através de aposentadoria ou outra atividade), PROD (faz manejo produtivo adequado<sup>18</sup>), ASSOC (*dummy* de participação em associações)

<sup>17</sup> Como explicado na descrição metodológica do modelo empírico, não seria possível tomar como adimplente um produtor que nunca contraiu crédito junto aos bancos públicos.

<sup>18</sup> Semelhante à variável PLANEJ, PROD leva em consideração respostas a questões referentes a práticas culturais que elevam a produtividade destas, quais sejam: realização de análises de solo e foliar, adoção das recomendações de análises laboratoriais, além de adoção de manejo específico para aumento do rendimento das culturas. As respostas possíveis seriam “não realiza”,

e SEMCONTR (*dummy* de venda principalmente através de transações sem qualquer tipo de contrato, mesmo informal).

Os resultados para os modelos específicos de melhor ajuste, tanto relativo às dívidas com o setor público, quanto às dívidas com o setor privado, são apresentados na tabela 5. Vale destacar que os campos vazios representam variáveis redundantes a cada tipo de modelo.

Novamente, as estatísticas LR —  $\chi^2(10)$  para o modelo de setor público e  $\chi^2(9)$  para o modelo de setor privado — e log pseudo-verossimilhança apontam para modelos apropriados. Para ambos os casos, todas as variáveis mostraram-se significativas para explicar a situação das dívidas dos produtores da amostra. Cabe destacar que, relativo ao modelo básico específico para o setor público, apenas as variáveis RENDA e IDADE apresentaram instabilidade nos efeitos marginais, o que pode ser justificado pela inclusão de variáveis correlacionadas no modelo de melhor ajuste. De forma análoga, para o modelo de setor privado, a instabilidade sobre os efeitos marginais também ocorre para a variável RENDA. Quanto ao número de observações utilizadas, tendo em vista a ocorrência de *missing values*, a amostra utilizada para o modelo de setor público corresponde a 715 lotes, enquanto equivale a 758 lotes no modelo de setor privado.

**Tabela 5:** Estatísticas dos modelos específicos de melhor ajuste

Discriminação	Dívida com Setor Público		Dívida com Setor Privado	
	Coefficientes	Ef. Marginais <sup>1</sup>	Coefficientes	Ef. Marginais <sup>1</sup>
AREA	0,04**	0,015**	-	-
RENDA	-0,09*	-0,035*	-0,05***	-0,006
ESCOL	-0,10*	-0,042*	-	-
IDADE	-0,15*	-0,058*	-0,09**	-0,010**
FILHOS	-	-	0,36**	0,031*
RENDEXTR	-	-	-0,18*	-0,020**
COLONO	0,27*	0,107*	-	-
NUMLOTES	0,08*	0,032*	-	-
CERTIF	-0,51*	-0,188*	-0,66*	-0,044*
PLANEJ	-0,08**	-0,033**	-0,07***	-0,008***
PROD	-	-	-0,09***	-0,010**
NUMFUNC	-0,02**	-0,010**	-	-
CONDFUNC	-0,58**	-0,212*	-	-
ASSOC	-	-	0,22**	0,029**
SEMCONTR	-	-	0,26*	0,036**
Elástica LR	$\chi^2(10) = 106,88^*$		$\chi^2(9) = 52,25^*$	
Log pseudo-verosim.	-677,58		-512,80	
Nº de observações	715		758	

\* Significativo a 1%. \*\* Significativo a 5%. \*\*\* Significativo a 10%.

Variável dependente: categoria 0 — adimplente; categoria 1 — renegociante; categoria 2 — inadimplente.

<sup>1</sup> Trata-se dos efeitos marginais de alterações em cada variável explanatória sobre as chances de ocorrência na categoria 2 (aqui apresentados apenas para verificação da estabilidade dos modelos). Fonte: elaboração própria.

Como observado, a hipótese de divergência dos fatores de influência e seus efeitos marginais quanto ao tipo de crédito é realmente verificada. Para expli-

“realiza parcialmente” ou “realiza”. Assim, da combinação desses resultados, a variável PROD foi criada variando de zero a quatro, de forma discreta e categórica, elevando-se conforme o grau de realização de manejo produtivo adequado.

car a inadimplência referente às dívidas contraídas junto ao setor público — levando-se em conta as instituições bancos públicos e CODEVASF —, contam positivamente os fatores AREA, COLONO e NUMLOTES, e negativamente as variáveis RENDA, ESCOL, IDADE, CERTIF, PLANEJ, NUMFUNC e CONDFUNC. Nesse caso, não são significativos ou são redundantes os fatores FILHOS, RENDEXTR, PROD, ASSOC e SEMCONTR. Por sua vez, para explicar a inadimplência às dívidas contraídas junto ao setor privado - levando-se em conta as instituições Distrito de Irrigação e casas comerciais -, contam positivamente os fatores FILHOS, ASSOC e SEMCONTR, e negativamente as variáveis RENDA, IDADE, RENDEXTR, CERTIF, PLANEJ e PROD. Nesse caso, não são significativos ou são redundantes os fatores AREA, ESCOL, COLONO, NUMLOTES, NUMFUNC e CONDFUNC.

Quanto às estimativas dos efeitos marginais, cabe salientar as características medianas do “produtor médio” para todas as variáveis explanatórias dispostas na tabela 5: AREA = 6,3 (área plantada no lote de 6,3 hectares); RENDA = 1 (renda bruta mensal do lote entre R\$ 2 mil e R\$ 3 mil); ESCOL = 4 (com Ensino Médio ou Técnico); IDADE = 3 (entre 50 e 60 anos); FILHOS = 1 (tem filhos dependentes); RENDEXTR = 1 (tem renda extra-rural, seja através de aposentadoria ou outra atividade); COLONO = 0 (não é caracterizado como colono); NUMLOTES = 1 (atua em apenas um lote); CERTIF = 0 (não certificado); PLANEJ = 3 (faz planejamento de negócio de forma parcial-forte<sup>19</sup>); PROD = 3 (faz manejo produtivo adequado de forma parcial-forte<sup>20</sup>); NUMFUNC = 1 (conta com um único funcionário); CONDFUNC = 0 (lote não conduzido por funcionário); ASSOC = 0 (não participa de associações); e SEMCONTR = 0 (apresenta, como principal forma de comercialização, transações por meio de contratos informais ou formais<sup>21</sup>). Estas seriam, portanto, as características do “produtor médio” do Polo Petrolina Juazeiro.

Antes de explanar sobre os efeitos marginais, é importante destacar que, de acordo com as estimativas realizadas a partir dos resultados da tabela 5, o “produtor médio” teria diferentes distribuições de probabilidades em cada categoria da variável dependente segundo o tipo de dívida, se contraída com o setor público ou com o setor privado. No caso do modelo para o setor público, a distribuição seria: chances de 16,98% de ser adimplente nas duas instituições tratadas (bancos comerciais e CODEVASF); 37,43% de renegociar dívidas com pelo menos uma das duas instituições; e 45,60% de se tornar inadimplente com pelo menos uma das duas instituições. Por sua vez, para o modelo de setor privado, a distribuição seria: chances de 83,44% de ser adimplente nas duas instituições tratadas (Distrito de Irrigação e casas comerciais); 10,88% de renegociar dívidas com pelo menos uma das duas instituições; e 5,68% de se tornar inadimplente com pelo menos uma das duas instituições.

Esse resultado, portanto, é tratado como a dimensão do fenômeno da inadimplência no Polo Petrolina-Juazeiro, qualificada para o tipo de dívida contraída. Como esperado, os produtores do Polo são mais comprometidos em pagar as dívidas contraídas com o setor privado do que são em pagar as dívidas contraídas junto a instituições públicas. A dimensão dessa divergência indica que as chances de o “produtor médio” ser inadimplente com o setor

<sup>19</sup> Como explanado anteriormente.

<sup>20</sup> Como explanado anteriormente.

<sup>21</sup> Contratos informais são tratados como aqueles casos em que o produtor exerce um contrato verbal ou, pelo menos, conhece o comprador devido a transações passadas. Contratos formais são aqueles em que o comprador firma a transação previamente por escrito.

público são 8,03 vezes maiores que as chances de o mesmo ser inadimplente com o setor privado. Por sua vez, as chances de ser renegociante nas dívidas com o setor público são 3,44 vezes maiores que as chances de o produtor ser renegociante com o setor privado. Por fim, as chances de inadimplência no setor privado são 4,91 vezes maiores que as chances de inadimplência no setor público.

Nesse sentido, o comprometimento ao pagamento das dívidas, por parte dos produtores do Polo, é discrepante segundo o tipo de instituição que se tem relação. A divergência apresentada é significativa, alarmante e reforça a hipótese de que pesa o histórico cultural de que as dívidas públicas serão perdoadas no futuro, o que configura um problema de principal-agente; e que as dívidas contraídas junto ao setor privado são caracterizadas como dívidas de curto prazo, o que leva o produtor a ser mais comprometido em efetuar seus pagamentos, sob o risco de impedir a continuidade de suas explorações produtivas.

Vale salientar que os resultados divergentes quanto à distribuição de probabilidades segundo o tipo de dívida entram em acordo com as observações amostrais a qual aponta para 50,3% de produtores da amostra como inadimplentes em uma das duas instituições do setor público, enquanto esse percentual é de 9,2% para o setor privado; 30,2% de renegociantes com o setor público e 15,5% com o setor privado; e 19,5% de inadimplentes com o setor público e 75,3% no caso das dívidas com as duas instituições aqui tratadas para o setor privado.

Passando finalmente aos resultados de efeitos marginais dos fatores determinantes da inadimplência, primeiro serão analisadas aquelas variáveis que são comuns a ambos os modelos, caso dos fatores RENDA, IDADE, CERTIF e PLANEJ. Quanto à variável RENDA, como se esperava, o aumento da renda bruta mensal do lote tende a reduzir as chances de inadimplência, visto que impacta diretamente sobre a capacidade de pagamento dos produtores. Os efeitos marginais referentes a esta variável, apresentados na tabela 6, indicam que, se o “produtor médio” atingir a faixa 6 (renda acima de R\$ 7 mil), suas chances de inadimplência são reduzidas de 45,60% para 29,02% no caso das dívidas com o setor público, enquanto são reduzidas de 5,68% para 3,35% no caso do setor privado.

**Tabela 6:** Efeitos marginais diante de choques na variável RENDA (modelos específicos)

Discriminação	Distribuição de probabilidades por estrato					
	Setor Público			Setor Privado		
	Adimpl.	Renegoc.	Inadimpl.	Adimpl.	Renegoc.	Inadimpl.
0: até R\$ 2 mil	14,84%	36,05%	49,12%	82,17%	11,56%	6,28%
1*: R\$ 2-3 mil	16,98%	37,43%	45,60%	83,44%	10,88%	5,68%
2: R\$ 3-4 mil	19,31%	38,58%	42,11%	84,65%	10,21%	5,13%
3: R\$ 4-5mil	21,82%	39,49%	38,69%	85,81%	9,56%	4,63%
4: R\$ 5-6 mil	24,52%	40,13%	35,35%	86,90%	8,94%	4,16%
5: R\$ 6-7 mil	27,38%	40,50%	32,12%	87,93%	8,33%	3,74%
6: acima de R\$ 7 mil	30,40%	40,58%	29,02%	88,91%	7,74%	3,35%

\* Valor referente ao “produtor médio”.

Fonte: Elaboração Própria

Com respeito à variável IDADE, produtores de maior faixa etária tendem a

ter maior compromisso de pagamento de suas dívidas. Dessa forma, conforme elucidada a tabela 7, se o “produtor médio” passasse a se caracterizar pela faixa 4 (acima de 60 anos), com tudo mais constante, suas chances de inadimplência seriam reduzidas para 39,82% no caso das dívidas com o setor público e para 4,75% no caso do setor privado.

**Tabela 7:** Efeitos marginais diante de choques na variável IDADE (modelos específicos)

Discriminação	Distribuição de probabilidades por estrato					
	Setor Público			Setor Privado		
	Adimpl.	Renegoc.	Inadimpl.	Adimpl.	Renegoc.	Inadimpl.
0: até 30 anos	8,11%	28,88%	63,00%	76,07%	14,57%	9,36%
1: 30-40 anos	10,56%	32,12%	57,32%	78,71%	13,32%	7,98%
2: 40-50 anos	13,51%	35,02%	51,47%	81,17%	12,08%	6,75%
3*: 40-60 anos	16,98%	37,43%	45,60%	83,44%	10,88%	5,68%
4: acima de 60 anos	20,97%	39,22%	39,82%	85,53%	9,72%	4,75%

\* Valor referente ao “produtor médio”.

Fonte: Elaboração Própria

Com relação à *dummy* CERTIF, a existência de certificação faz as chances de o “produtor médio” se tornar inadimplente reduzir consideravelmente, tanto no caso das dívidas contraídas no setor público, quanto no privado. Nesse sentido, no modelo de setor público, a distribuição de probabilidades de um “produtor médio” agora certificado passaria a ser de 32,71% de ser adimplente, 40,45% de se tornar renegociante e 26,83% de ser inadimplente. Por sua vez, no caso do modelo de setor privado, a distribuição passaria a ser de 94,86% de se tornar adimplente, 3,89% de ser renegociante e 1,25% de ser inadimplente.

Quanto à variável PLANEJ, observa-se que a realização de práticas de gestão financeira no negócio tende a reduzir as chances de inadimplência, como apresentado na tabela 8. Por sua vez, se o “produtor médio” não realiza qualquer gestão financeira, mesmo que parcialmente, suas chances de inadimplência são elevadas para 55,52% no caso de dívidas contraídas junto ao setor público e para 4,92% no caso de dívidas no setor privado.

**Tabela 8:** Efeitos marginais diante de choques na variável PLANEJ (modelos específicos)

Discriminação	Distribuição de probabilidades por estrato					
	Setor Público			Setor Privado		
	Adimpl.	Renegoc.	Inadimpl.	Adimpl.	Renegoc.	Inadimpl.
0: Não realiza	11,42%	33,06%	55,52%	77,66%	13,82%	8,52%
1: Parcial-fraco	13,11%	34,67%	52,22%	79,70%	12,82%	7,47%
2: Parcial-médio I	14,96%	36,14%	48,91%	81,63%	11,84%	6,53%
3*: Parcial-forte	16,98%	37,43%	45,60%	83,44%	10,88%	5,68%
4: Realiza totalmente	19,16%	38,52%	42,32%	85,13%	9,95%	4,92%

\* Valor referente ao “produtor médio”.

Fonte: Elaboração Própria

Passando para a análise dos efeitos marginais dos fatores determinantes da inadimplência unicamente relacionados às dívidas contraídas com o setor público, a variável AREA indica que um aumento de 1 ha sobre a área plantada

no lote do “produtor médio” (levando-o a se caracterizar por AREA = 7,6) acarretaria uma pequena alteração na distribuição de probabilidades, conforme demonstra a tabela 9. Contudo, o aumento de 10 ha, por exemplo, seria capaz de elevar suas chances de inadimplência de 45,60% para 61,03%, enquanto reduziria as chances de adimplência para apenas 8,92%. Dessa forma, se o “produtor médio” passar a ter um maior porte, mantendo-se constantes todas as demais características, suas chances de inadimplência às dívidas no setor público tendem a ser elevadas fortemente.

Como comentado anteriormente, vale destacar que a variável AREA mostrou-se redundante no modelo explicativo da situação das dívidas dos produtores da amostra relativas ao setor privado.

**Tabela 9:** Efeitos marginais diante de choques na variável AREA (setor público)

Discriminação	Distribuição de probabilidades por estrato		
	Adimplente	Renegociante	Inadimplente
Mínimo amostral (0,2 ha)	23,67%	39,97%	36,36%
“Produtor médio” (6,3 ha)	16,98%	37,43%	45,60%
1 ha + “produtor médio” (7,3 ha)	16,01%	36,84%	47,15%
10 ha + “produtor médio” (16,3 ha)	8,92%	30,05%	61,03%
Máximo amostral (28 ha)	3,57%	19,48%	76,95%

Fonte: Elaboração própria.

Com respeito à variável ESCOL, verifica-se que uma maior graduação escolar tende a reduzir as chances de inadimplência às dívidas contraídas junto ao setor público, ainda que de forma discreta — se o “produtor médio” passar a ter Ensino Superior, as chances de adimplência crescem para apenas 19,76%, conforme apresentado na tabela 10. Para o setor privado, por sua vez, a variável ESCOL mostrou-se redundante para explicar a situação da dívida.

**Tabela 10:** Efeitos marginais diante de choques na variável ESCOL (setor público)

Discriminação	Distribuição de probabilidades por estrato		
	Adimplente	Renegociante	Inadimplente
0: Analfabeto	8,47%	29,42%	62,11%
1: Alfabetizado	10,22%	31,72%	58,07%
2: Ensino Fundamental I	12,21%	33,85%	53,94%
3: Ensino Fundamental II	14,46%	35,77%	49,77%
4*: Ensino Médio ou Técnico	16,98%	37,43%	45,60%
5: Ensino Superior	19,76%	38,77%	41,48%

\* Valor referente ao “produtor médio”.

Fonte: Elaboração Própria

Com relação à *dummy* COLONO, o produtor caracterizado como colono, ou seja, atuando no lote como meeiro, arrendatário ou em regime de comodatado, tende a ter maiores chances de ser inadimplente em uma das duas instituições aqui tratadas no âmbito do setor público. Caso o “produtor médio” passasse a ser caracterizado como colono (COLONO = 1), sua distribuição de probabilidades seria alterada para 11,02% para adimplência, 32,64% para renegociação e 56,34% para inadimplência frente às dívidas contraídas junto ao

setor público. Para o setor privado, contudo, a *dummy* COLONO mostrou-se redundante para explicar a variável dependente.

O número de lotes em que o produtor atua também influencia a situação da dívida do mesmo frente ao setor público, diferente do que ocorre para o setor privado, como explanado anteriormente, em que a variável NUMLOTES mostrou-se redundante para explicar a inadimplência. Assim, se o “produtor médio” passar a atuar em dois lotes, as chances de inadimplência no setor público são elevadas para 48,80%. No caso de o mesmo assumir o valor máximo amostral, de atuação em dez lotes, referida probabilidade cresce para 73,01%, conforme apresentado na tabela 11.

**Tabela 11:** Efeitos marginais diante de choques na variável NUMLOTES (setor público)

Discriminação	Distribuição de probabilidades por estrato		
	Adimplente	Renegociante	Inadimplente
“Produtor médio” (1 lote)	16,98%	37,43%	45,60%
1 + “produtor médio” (2 lotes)	15,02%	36,18%	48,80%
Máximo amostral (10 lotes)	4,66%	22,33%	73,01%

Fonte: Elaboração própria.

Com respeito à variável NUMFUNC, quanto maior o número de funcionários que trabalha no lote, menores as chances de o produtor se tornar inadimplente frente às dívidas relativas ao setor público, conforme apresenta a tabela 12. Caso o “produtor médio” tenha mais um funcionário, as chances de inadimplência caem para 44,61%. Apesar da variação discreta, esse é um indicativo de que produtores caracterizados pela agricultura familiar tendem a ter maiores chances de descumprimento às dívidas tomadas no setor público. Como explicado anteriormente, tal sentença não se aplica ao caso das dívidas junto ao setor privado, cuja variável NUMFUNC mostrou-se redundante para explicar a variável dependente.

**Tabela 12:** .

Discriminação	Distribuição de probabilidades por estrato		
	Adimplente	Renegociante	Inadimplente
Mínimo amostral (nenhum funcionário)	16,35%	37,06%	46,59%
“Produtor médio” (1 funcionário)	16,98%	37,43%	45,60%
1 + “Produtor médio” (2 funcionários)	17,61%	37,77%	44,61%
Máximo amostral (56 funcionários)	66,05%	27,00%	6,95%

\* Valor referente ao “produtor médio”.

Fonte: Elaboração Própria

Por fim, quanto à *dummy* CONDFUNC, quando o “produtor médio” caracteriza-se por ter a administração de seu lote conduzido por funcionário, sua distribuição de probabilidades seria de 35,44% de chances de inadimplência, 40,12% de chances de renegociação e 24,44% de ser inadimplente quanto às dívidas contraídas junto ao setor público. Nesse sentido, a gestão profissional do negócio, no caso da amostra aqui analisada, tende a elevar as chances de inadimplência do produtor. Contudo, para as dívidas no setor privado, a *dummy* CONDFUNC mostra-se redundante para explicar a inadimplência.

Passando para a análise dos efeitos marginais dos fatores determinantes da inadimplência unicamente relacionados às dívidas contraídas com o setor privado, a *dummy* FILHOS, referente à existência de filhos dependentes, indica uma relação positiva com a inadimplência nesse tipo de dívida. Assim, se o “produtor médio” não possuísse filhos dependentes, sua distribuição de probabilidades seria de 90,85% de chances de adimplência, 6,54% de chances de renegociação e 2,61% de ser inadimplente. Portanto, a inexistência de filhos dependentes tende a reduzir a probabilidade de o produtor não cumprir suas obrigações frente às dívidas contraídas junto ao setor privado. Quanto ao setor público, contudo, referida *dummy* é redundante para explicar a variável dependente.

Por sua vez, quanto à *dummy* RENDEXTR, o “produtor médio” caracteriza-se por possuir algum tipo de renda extrarrural, seja através de aposentadoria ou através de outra atividade. Caso este não possuísse renda extrarrural, sua distribuição de probabilidades seria de 78,66% de chances de adimplência, 13,34% de chances de renegociação e 8,00% de ser inadimplente quanto às dívidas contraídas junto ao setor privado. Quanto ao setor público, assim como descrito para a *dummy* FILHOS, a disponibilidade de renda extra-rural é redundante para explicar a inadimplência.

Com respeito à variável PROD, esta tende a reduzir as chances de inadimplência no setor privado, diferente do que ocorre no caso das dívidas contraídas no setor público, cuja variável é redundante para explicar a categoria em que o produtor se encontra. Assim, observa-se que a realização de práticas produtivas adequadas tende a reduzir as chances de inadimplência no setor privado, como apresentado na tabela 13. Por sua vez, se o “produtor médio” não realiza qualquer manejo produtivo, mesmo que parcialmente, suas chances de inadimplência são elevadas para 9,29%.

**Tabela 13:** Efeitos marginais diante de choques na variável PROD (setor privado)

Discriminação	Distribuição de probabilidades por estrato		
	Adimplente	Renegociante	Inadimplente
0: Não realiza manejo adequado	76,20%	14,51%	9,29%
1: Realiza / parcialmente-fraco	78,79%	13,28%	7,94%
2: Realiza / parcialmente-médio I	81,20%	12,06%	6,74%
3*: Realiza / parcialmente-forte	83,44%	10,88%	5,68%
4: Realiza completamente	85,50%	9,74%	4,76%

\* Valor referente ao “produtor médio”.

Fonte: Elaboração Própria

Quanto à *dummy* ASSOC, quando o “produtor médio” caracteriza-se por participar de associação de produtores, sua distribuição de probabilidades seria de 77,50% de chances de adimplência, 13,90% de chances de renegociação e 8,60% de ser inadimplente quanto às dívidas contraídas junto ao setor privado. Nesse sentido, a participação em associações tende a elevar as chances de inadimplência do produtor<sup>22</sup>. Contudo, para as dívidas no setor público, a *dummy* ASSOC mostra-se redundante para explicar a inadimplência.

<sup>22</sup>Tal resultado pode estar associado a fatores como pouca atuação e desvalorização das organizações de produtores no Polo.

Por fim, com respeito à *dummy* SEMCONTR, o “produtor médio” apresenta, como principal forma de comercialização, transações por meio de contratos informais ou formais. Assim, se o “produtor médio” efetuasse suas vendas por meio de estruturas menos coordenadas de transação, sua distribuição de probabilidades seria de 76,19% de chances de adimplência, 14,51% de chances de renegociação e 9,29% de ser inadimplente quanto às dívidas contraídas junto ao setor privado. Para as dívidas no setor público, referida variável mostra-se redundante para explicar a variável dependente.

Ademais, cabe ressaltar que a análise dos efeitos marginais foi feita individualmente para cada variável explanatória, segundo cada modelo estimado (conjunto e específicos). Em todas as análises, os choques provocados foram individuais, mantendo-se constantes as demais variáveis. Contudo, é óbvio que a distribuição de probabilidades poderá ser alterada diante de choques conjuntos. Dessa forma, foram aqui apresentados modelos que permitem a estimativa das chances de produtores com características variadas serem ou não adimplentes, renegociantes ou inadimplentes, configurando-se, portanto, em importante ferramenta para se levantar hipóteses sobre o risco de crédito relativo a cada produtor.

Por exemplo, considere-se um produtor cujas características são favoráveis e descritas como: AREA = 6,3 (equivalente à área plantada no lote do “produtor médio”); RENDA = 6 (renda mensal do lote acima de R\$ 7 mil); ESCOL = 5 (com Ensino Superior); IDADE = 3 (entre 50 e 60 anos, equivalente ao “produtor médio”); FILHOS = 1 (tem filhos dependentes, como o “produtor médio”); RENDEXTR = 1 (tem renda extra-rural, equivalente ao “produtor médio”); COLONO = 0 (não é caracterizado como colono, assim como o “produtor médio”); NUMLOTES = 1 (atua em apenas um lote, assim como o “produtor médio”); CERTIF = 1 (tem certificação); PLANEJ = 4 (faz gerenciamento financeiro); PROD = 4 (faz manejo produtivo adequado); NUMFUNC = 6 (conta com seis funcionários - um por hectare plantado); CONDFUNC = 0 (lote não conduzido por funcionário, equivalente ao “produtor médio”); ASSOC = 0 (não participa de associações, equivalente à área do “produtor médio”); e SEMCONTR = 0 (transaciona por meio de contratos informais e/ou formais, assim como o “produtor médio”).

Como contraponto, considere-se outro exemplo, de um produtor cujas características são desfavoráveis e descritas como: AREA = 6,3 (equivalente à área plantada no lote do “produtor médio”); RENDA = 0 (renda bruta mensal do lote inferior a R\$ 2 mil); ESCOL = 5 (com Ensino Médio ou Técnico, assim como o “produtor médio”); IDADE = 3 (entre 50 e 60 anos, assim como o “produtor médio”); FILHOS = 1 (tem filhos dependentes, assim como o “produtor médio”); RENDEXTR = 0 (não tem renda extra-rural); COLONO = 0 (não é caracterizado como colono, equivalente ao “produtor médio”); NUMLOTES = 1 (atua em apenas um lote, equivalente ao “produtor médio”); CERTIF = 0 (não tem certificação, equivalente ao “produtor médio”); PLANEJ = 0 (não faz qualquer gerenciamento financeiro, mesmo que parcialmente); PROD = 0 (não faz manejo produtivo adequado, mesmo que parcialmente); NUMFUNC = 1 (conta com um funcionário, assim como o “produtor médio”); CONDFUNC = 0 (lote não conduzido por funcionário, assim como o “produtor médio”); ASSOC = 0 (não participa de associações, assim como o “produtor médio”); e SEMCONTR = 1 (não transaciona por meio de contratos, mesmo informais).

No caso do produtor com condições favoráveis, as chances de adimplência estimadas seriam de 54,73% no modelo conjunto, 62,04% no modelo espe-

cífico referente ao setor público e 97,92% no modelo específico referente ao setor privado. Por outro lado, no caso do produtor com condições desfavoráveis, as chances de adimplência estimadas seriam de apenas 7,15% no modelo conjunto, 9,80% no modelo específico referente ao setor público e 50,62% no modelo específico referente ao setor privado.

### 4.3 Ajustes dos Modelos

Para efeitos de exercício para uma medida de ajuste dos modelos aos dados amostrais, foram calculadas as probabilidades de ocorrência em cada uma das categorias da variável dependente em cada modelo estimado — dadas as informações observadas para o vetor  $X$  —, sendo esses resultados confrontados com o valor realmente observado da variável dependente. Para tanto, foram considerados como erros as seguintes ocorrências: i) indicar ser mais provável ser inadimplente quando, na verdade, o produtor foi inadimplente ou renegociante; ii) indicar ser mais provável ser inadimplente quando, na verdade, o produtor foi adimplente ou renegociante; e iii) indicar ser mais provável ser renegociante quando, na verdade, o produtor foi adimplente ou inadimplente. Dessa maneira, verificou-se que houve 56,45% de acertos no caso do modelo conjunto (dentre as 659 observações utilizadas nesse modelo), 56,22% no modelo específico para o setor público (dentre as 715 observações utilizadas) e 76,12% no modelo específico para o setor privado (dentre as 758 utilizadas).

Obviamente, ao se relaxar as condições de erro propostas, os resultados são ainda mais satisfatórios. Assim, resolveu-se dar uma medida de ajuste apenas sobre os casos extremos, de adimplência e inadimplência. Nesse caso, verificou-se que:

- Quanto ao modelo conjunto: i) se as estimativas indicaram maior probabilidade de que o produtor fosse adimplente nas quatro instituições consideradas, em 83,3% desses casos este não foi inadimplente em nenhuma delas; e ii) se as estimativas indicaram maior probabilidade de que o produtor fosse inadimplente em uma das quatro instituições consideradas, em 90,5% desses casos este não foi adimplente em todas elas;
- Quanto ao modelo específico para o setor público: i) se as estimativas indicaram maior probabilidade de que o produtor fosse adimplente nas duas instituições consideradas, em 82,7% desses casos este não foi inadimplente em nenhuma delas; ii) se as estimativas indicaram maior probabilidade de que o produtor fosse inadimplente em uma das duas instituições consideradas, em 87,1% desses casos este não foi adimplente em todas elas;
- Quanto ao modelo específico para o setor privado: em todas as observações, as estimativas indicaram maior probabilidade de que o produtor fosse adimplente nas duas instituições consideradas e, assim, em 90,1% desses casos este não foi inadimplente em nenhuma delas.

Tais resultados indicam um ajuste satisfatório para o levantamento de hipóteses sobre o risco de crédito relativo a cada produtor.

## 5 Considerações Finais

Diante dos resultados apresentados, verifica-se que a dimensão da inadimplência no Polo Petrolina-Juazeiro é preocupante, no sentido de que as chances de um “produtor médio” do Polo se tornar inadimplente em pelo menos uma das quatro instituições aqui tratadas (bancos públicos, CODEVASF, Distrito de Irrigação e casas comerciais) chega a 46,09%. Do lado oposto, a probabilidade de adimplência em todas as quatro instituições chega a somente 14,64%, enquanto as chances de renegociação chegam a 39,26%.

Contudo, verificou-se que o descompromisso com as dívidas por parte dos produtores do Polo varia de acordo com o tipo de instituição. Nesse sentido, verifica-se que o problema da inadimplência no Polo está fortemente relacionado com as dívidas contraídas junto a instituições do setor público, aqui tratadas como os bancos públicos e a CODEVASF. Para o caso das dívidas com o setor público, a estimativa das chances de inadimplência chega a 45,60%, cerca de 8 vezes mais do que as chances de inadimplência no setor privado — aqui representado pelo Distrito de Irrigação e pelas casas comerciais. Do lado oposto, a probabilidade de adimplência no setor privado chega a 83,44%, enquanto que no setor público essas chances são de apenas 16,98%. Na situação intermediária, de renegociação das dívidas, as chances no setor público foram estimadas em 37,43%, valor 3,4 vezes maior que as chances de renegociação com uma das duas instituições aqui tratadas como sendo do setor privado.

Além disso, os fatores determinantes da situação das dívidas são peculiares a cada tipo de endividamento. No caso da inadimplência às dívidas com o setor público, pesam positivamente a renda bruta do lote, a escolaridade e a idade do produtor, a existência de certificação (Globalgap), a realização de práticas de gerenciamento financeiro, o número de funcionários no lote e produtores com funcionários que conduzem o lote; e negativamente a área plantada no lote, a condição de colono (meeiro, arrendatário ou em regime de comodato) e o número de lotes em que o produtor atua. Por outro lado, para a inadimplência às dívidas com o setor privado, pesam positivamente a renda bruta do lote, a idade do produtor, a existência de renda extra-rural, a existência de certificação (Globalgap), a realização de práticas de gerenciamento financeiro e manejo produtivo adequado; e negativamente a condição de ter filhos dependentes, participar de associações e não vender suas frutas por meio de contratos informais e/ou formais.

Os resultados encontrados neste trabalho concatenam com os encontrados por autores como Antão & Campanholo (2011), Villa Verde (2000) e Costa (2007), onde se mostra que existe uma tendência no sentido dos produtores agrícolas serem menos comprometidos em pagar as dívidas junto ao setor público do que em relação ao setor privado. Isso se deve sobretudo ao histórico cultural de que as dívidas públicas seriam perdoadas no futuro e também por essas dívidas se caracterizarem como dívidas de longo prazo (menos comprometimento em efetuar seus pagamentos). As dívidas contraídas junto ao setor privado, ao contrário, seriam dívidas de curto prazo, o que afetaria diretamente as atividades de exploração produtiva e, conseqüentemente, sua continuidade.

Vale destacar que a análise do modelo conjunto aponta para a situação geral da inadimplência no Polo, sem qualificar o tipo de instituição em que as dívidas são contraídas. Por esse motivo, seus resultados concatenam os fenômenos específicos da inadimplência frente ao setor público e frente ao setor

privado. Ademais, os diagnósticos de ajuste dos modelos estimados (conjunto e específicos) apontam para níveis satisfatórios que permitem a utilização dos mesmos para inferir sobre o risco de crédito associado a determinado produtor que se deseja estudar, diante de suas características.

Ante ao acima exposto, o foco de políticas setoriais deve recair sobre os seguintes aspectos: i) incentivar a realização de práticas de gerenciamento financeiro entre os produtores; ii) promover o acesso dos produtores a técnicas produtivas que elevem a produtividade dos cultivos explorados; iii) estimular a adequação dos produtores aos níveis de tecnologia e qualidade exigidos internacionalmente; iv) efetuar medidas de redução da incidência do comércio realizado por meio de atravessadores, estimulando a realização de contratos entre produtores e demandantes (garantia de preço); v) reformular a gestão das organizações de produtores locais, dando maior dinamismo às estratégias de associativismo; vi) dar especial atenção aos colonos, estimulando práticas produtivas e gerenciais adequadas; e vii) promover medidas voltadas à melhoria da qualidade de ensino na região, principalmente do ensino técnico aplicado à agricultura, com potenciais efeitos de longo prazo. Para que a problemática atual da inadimplência seja contornada, é necessário que medidas nesse sentido sejam colocadas em prática, com o objetivo de se buscar, de maneira responsável e construtiva, uma solução racional e justa, com a participação e envolvimento dos produtores e seus agentes relacionados, dos bancos, da sociedade e do Governo.

## Referências Bibliográficas

- Akerlof, G. (1970), 'The market for lemons: Qualitative uncertainty and the market mechanism', *Quarterly Journal of Economics* 3(84), 488–500.
- Antão, R. A. S. & Campanholo, T. (2011), 'O crédito rural no contexto do desenvolvimento econômico e social', *Revista da Católica* 3(5), 1–12.
- Arrow, K. J. (1971), *Essays in the Theory of Risk-Bearing*, Markham Publishing Co.
- Azevedo, C. M. & Shikida, P. F. A. (2004), 'Assimetria de informação e o crédito agropecuário: o caso dos cooperados da coamo-toledo (pr)', *Revista de Economia e Sociologia Rural* 42(2), 267–292.
- Brant, R. (1990), 'Assessing proportionality in the proportional odds model for ordinal logistic regression', *Biometrics* 46(4), 1171–8.
- Caouette, J. B., Altman, E. I. & Narayanan, P. (1999), *Gestão do Risco de Crédito*, Qualitymark Editora Ltda.
- Cardoso, L. V. (2005), *Coordenação e competitividade no sistema agroindustrial de mamão: estudo multicaso de firmas exportadoras*, Dissertação (mestrado em agronegócios), UFMS/UFGO/UNB, Brasília-DF.
- Chaddad, F. R. & Lazzarini, S. G. (2003), 'Relações contratuais de crédito agrícola e o papel dos agentes financeiros privados: teoria e evidências dos eua', *Revista de Economia e Sociologia Rural* 41(3), 29–52.

CODEVASF (2012), 'Companhia de desenvolvimento dos vales do são francisco e do parnaíba'. <http://www.codevasf.gov.br/>. Acessado em junho de 2012.

Costa, J. G. (2007), Entraves e alternativas para a expansão do crédito rural na fruticultura irrigada do polo petrolina-juazeiro, Dissertação (mestrado em economia), UFPE, Recife-PE.

Costa, S. M. A., Gomes, M. R. L. & Tarsitano, M. A. A. (2008), 'A comercialização de uvas finas na região de jales-sp', *Revista Brasileira de Fruticultura* 30(1), 127–132.

Davidson, R. & MacKinnon, J. (2003), *Econometric theory and methods*, 1 ed. edn, Oxford University Press.

Dias, G. L. S. (1991), 'Mobilização de poupança e o financiamento do setor agrícola', Piracicaba: FEALQ/CEPEA.

Faria, R. N., Souza, R. C., Vieira, J. G. V. & Lírio, V. S. (2008), 'Custo de transação e exigências técnicas nas exportações de manga e de mamão', *Informações Econômicas* 38(5), 59–71.

Franca, C. S. (2004), Impacto da globalização e modernização agrícola na região do vale submédio do são francisco: estudo de caso do perímetro de irrigação projeto bebedouro em petrolina - pe, Dissertação (mestrado em sociologia), UFPE, Recife-PE.

Gatti, E. U., Vieira, J. L. T. M. & Silva, V. (1993), 'Análise do perfil distributivo do crédito rural no brasil, 1969-90', *Agricultura em São Paulo* 40(1), 65–99.

Greene, W. H. (2003), *Econometric Analysis*, 5 ed. edn, Prentice Hall.

Guanziroli, C. E. (2007), 'Pronaf dez anos depois: resultados e perspectivas para o desenvolvimento rural', *Revista de Economia e Sociologia Rural* 45(2), 301–328.

Lazzarini, S. G., Neves, M. F. & Chaddad, F. R. (1999), 'Pacotes de insumos agrícolas: Mudando a configuração do agribusiness', *Preços Agrícolas* 13(154), 52–53.

Ortega, A. C. & Só, L. S. (2005), 'Impacto das políticas do mda: um estudo de caso do pronaf nos municípios de orizona e davinópolis em goiás (brasil)', *Anais...*, Santiago, Chile.

PRONAF (2006), 'Programa nacional de fortalecimento da agricultura familiar'. <http://portal.mda.gov.br/portal/saf/programas/pronaf>. Acessado em junho de 2014.

Rocha Júnior, W. F. (2001), A matriz estrutural prospectiva com o enfoque da nova economia institucional: o caso do agronegócio da erva-mate, Tese (doutorado em economia), UFSC, Florianópolis-SC.

Sherrick, B. J., Sonka, S. & Monke, J. D. (1994), 'Nontraditional lenders in agricultural credit markets', *Agribusiness* 10(4), 341–357.

Sobel, T. F. (2011), Fruticultura e Economia dos Custos de Transação: determinantes das estruturas de governança dos pequenos produtores do Polo Petrolina-Juazeiro, Tese (doutorado em economia), UFPE, Recife-PE.

STATA (2012), 'Data analysis and statistical software'. <http://www.stata.com/>. Acessado em junho de 2012.

Staten, M. E., Gilley, O. W. & Umbeck, J. (1990), 'Information costs and the organization of credit markets: A theory of indirect lending', *Economic Inquiry* 28(3), 508–29.

Stiglitz, J. E. (2000), 'The contributions of the economics of information to twentieth century economics', *Quarterly Journal of Economics* 115(4), 1441–1478.

Stiglitz, J. E. & Weiss, A. (1981), 'Credit rationing in markets with imperfect information', *American Economic Review* 71(3), 393–410.

Trombin, V. G. (2007), Proposição de um método para analisar a viabilidade da implantação de uma cadeia produtiva em um novo local: o caso da citricultura no pólo petrolina juazeiro, Dissertação (mestrado em administração de organizações), USP, São Paulo-SP.

Villa Verde, C. M. (2000), 'O crédito rural e a capacidade de pagamento do setor agrícola', IPEA, Brasília.

Wang, X. & Kockelman, K. M. (2005), 'Occupant injury severity using a heteroscedastic ordered logit model: distinguishing the effects of vehicle weight and type', *Annals of...*, Washington.

Williamson, O. E. (1995), *The economic institutions of capitalism*, The Free Press.

Wolfe, F. & Gould, W. (1998), 'An approximate likelihood-ratio test for ordinal response models', *Stata Technical Bulletin* 7(42), 24–27.

Zylbersztajn, D. (2005), 'Papel dos contratos na coordenação agro-industrial: um olhar além dos mercados', *Revista de Economia e Sociologia Rural* 43(3), 385–420.

## Informações aos autores

### Objetivo e escopo

A ECONOMIA APLICADA é uma publicação trimestral do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, com uma linha editorial voltada exclusivamente para artigos científicos em economia aplicada. Tem por foco a publicação de trabalhos de análise econômica aplicada a problemas específicos, tanto de interesse do setor público como privado, e que contenham, particularmente, estudos quantitativos, cujos resultados possam aproximar a teoria da realidade. A interação entre trabalho empírico e teoria, análise e política econômica é, portanto, um importante diferencial da revista. A revista ECONOMIA APLICADA coloca-se como um espaço aberto para toda a comunidade acadêmica, nacional e internacional.

### Política editorial

Os artigos submetidos são avaliados inicialmente pelos Editores quanto a sua adequação à linha editorial da Revista. Uma vez aceita a submissão, o artigo é encaminhado para avaliação por pareceristas designados pelos Editores, no sistema double-blind.

### Apresentação de Artigos

Os artigos submetidos para avaliação devem obedecer às seguintes normas:

1. Apresentar ineditismos no país e no exterior
2. Na primeira página devem constar o título do artigo, nome dos autores com endereço para contato e afiliação institucional, resumo com até 100 palavras em português e inglês, palavras chave, em português e inglês e classificação JEL.
3. Na versão para “avaliação”, basta um arquivo no formato pdf com todas as equações, tabelas e figuras no corpo do texto, e referências no final.
4. Caso o artigo venha a ser aceito para publicação, a equipe da revista contatará os editores para as atividades de editoração. Nessa ocasião serão requeridos os seguintes arquivos:
  - Arquivo fonte no formato  $\LaTeX$ .
  - Arquivo de referências bibliográficas no padrão Bib $\TeX$ .
  - Arquivo com as figuras em formato TIFF ou EPS.
  - Planilha eletrônica com todas as tabelas.

## Responsabilidade e Direitos Autorais

Os autores que submeterem ou tiverem o artigo aceito para publicação na Revista ECONOMIA APLICADA declaram:

- Ter participado da concepção do artigo para tornarem públicas suas responsabilidades pelos seus conteúdos, que não omitiram quaisquer ligações ou acordos de financiamento entre autores e companhias que possam ter interesse na publicação deste artigo.
- Que o manuscrito é inédito e que o artigo, em parte ou na íntegra, ou qualquer outro artigo com conteúdo substancialmente similar, de sua autoria, não foi enviado à outra revista e não o será enquanto sua publicação estiver sendo considerada pela revista ECONOMIA APLICADA, que seja no formato impresso ou no eletrônico, exceto o descrito em anexo.
- Que a revista ECONOMIA APLICADA passa a ter os direitos autorais para a veiculação dos artigos, tanto em formato impresso como eletrônico a eles referentes, os quais se tornarão propriedade exclusiva da ECONOMIA APLICADA. É permitida a reprodução total ou parcial dos artigos desta revista, desde que seja explicitamente citada a fonte completa.

ISSN 1413-8050



The Brazilian Journal of Applied Economics is a quarterly publication of the Departments of Economics of the FEA-RP of the University of São Paulo. The published papers are sole responsibility of the author(s). As the author(s) send a manuscript for submission to the Journal they tacitly agree that, if accepted for publication, the copyright of the material will be transferred to the ECONOMIA APLICADA (BRAZILIAN JOURNAL OF APPLIED ECONOMICS). The Journal may then publish the material in paper and electronic formats that will be exclusive property of the ECONOMIA APLICADA (BRAZILIAN JOURNAL OF APPLIED ECONOMICS). The publication of the whole papers or parts of it is permitted as long as the source is explicitly mentioned. The Brazilian Journal of Applied Economics is published with the financial support of the CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico. This journal is indexed by Journal of Economic Literature, electronic on line and CD-ROM, and Scielo Brasil.

Secretary, Sales and Delivery: Juliana de Freitas Vitória      Cover: Christof Gunkel  
Linguistic revision (Portuguese and English): Iulo Feliciano Afonso (Solution Translations Eireli — Epp)  
Diagramming: Francisco Nunes Navarro & Bruna Carolina Garcia

### **Mailing Address**

Revista Economia Aplicada, Departamento de Economia, FEA-PR/USP  
Avenida do Bandeirantes, 3900, Ribeirão Preto, SP.  
CEP 14040-905  
Telephone: (0xx16) 3602-3910, fax: (0xx16) 3633-4488  
email:revecap@usp.br  
site: <http://www.usp.br/revecap>

### **Subscriptions for Delivery Outside Brazil**

Individuals    \$ 90.00  
Institutions   \$110.00

\* Air mail included

† Individual issues can  
also be purchased at  
the above address.

## Editor

Reynaldo Fernandes

## Associate Editors

Luiz Guilherme Scorzafave	Sergio Naruhiko Sakurai
Angelo Costa Gurgel	Eliezer Martins Diniz
Rodrigo De Losso	Irineu E. de Carvalho Filho
Cláudio Ribeiro de Lucinda	Walter Belluzzo Junior

## Editorial Board

<b>Affonso Celso Pastore</b> USP	<b>Antônio Barros de Castro</b> UFRJ
<b>Cássio F. Camargo Rolim</b> UFPR	<b>Cláudio Monteiro Considera</b> UFF
<b>Clélio Campolina Diniz</b> CEDEPLAR	<b>Denisard C. de Oliveira Alves</b> USP
<b>Eleutério F. S. Prado</b> USP	<b>Fernando de Holanda Barbosa</b> FGV-UFF
<b>Francisco Anuatti Neto</b> USP/RP	<b>Geoffrey J. D. Hewings</b> University of Illinois
<b>Geraldo Sant'ana de Camargo Barros</b> ESALQ/USP	<b>Gustavo Maia Gomes</b> IPEA
<b>José Marcelino da Costa</b> NAEA/PA	<b>José A. Scheinkman</b> Princeton University
<b>Marcelo Portugal</b> UFRGS	<b>Maria José Willumsen</b> Florida International University
<b>Márcio Gomes Pinto Garcia</b> PUC/ RJ	<b>Mário Luiz Possas</b> UFRJ
<b>Paulo César Coutinho</b> UnB	<b>Paulo Nogueira Batista Júnior</b> FGV/SP
<b>Pierre Perron</b> Boston University	<b>Pedro Cesar Dutra Fonseca</b> UFRGS
<b>Ricardo R. Araújo Lima</b> UnB	<b>Robert E. Evenson</b> Yale University
<b>Roberto Smith</b> UFCE	<b>Rodolfo Hoffmann</b> ESALQ/USP
<b>Rogério Studart</b> UFRJ	<b>Russell E. Smith</b> Washburn University
<b>Sérgio Werlang</b> FGV/RJ	<b>Tomás Málaga</b> FGV/SP
<b>Victor Bulmer-Thomas</b> University of London	<b>Werner Baer</b> University of Illinois
<b>Wilson Suzigan</b> Unicamp	