

**RISCO DAS INSTITUIÇÕES FINANCEIRAS E O IMPACTO SOBRE A
ATIVIDADE ECONÔMICA: UMA ANÁLISE COM DADOS EM
PAINEL PARA O BRASIL**

Helder Ferreira de Mendonça
Délio José Cordeiro Galvão
Renato Falci Villela Loures

**DIREITOS DE PROPRIEDADE E BEM-ESTAR: AVALIAÇÃO DO IMPACTO
DO PROGRAMA DE REGULARIZAÇÃO FUNDIÁRIA
NA QUINTA DO CAJU**

Maria Isabel de Toledo Andrade
Valéria Pero

**TRABALHO MATERNO E DESEMPENHO EDUCACIONAL DAS CRIANÇAS:
UMA ANÁLISE DA PROBABILIDADE DE APROVAÇÃO ESCOLAR**

Juliana Maria de Aquino
Elaine Toldo Pazello

**GASTO EM P&D E PODER DE MERCADO: TEORIA E EVIDÊNCIA
PARA O BRASIL**

Eduardo Correia
Rodrigo Moita

**UM MODELO ECONOMÉTRICO COM PARÂMETROS VARIÁVEIS
PARA A CARGA TRIBUTÁRIA BRUTA BRASILEIRA TRIMESTRAL**

Mario Jorge Cardoso de Mendonça
Adolfo Sachsida
Luis Alberto Medrano

Governo Federal

Secretaria de Assuntos Estratégicos da Presidência da República

Ministro – Wellington Moreira Franco



**Instituto de Pesquisa
Econômica Aplicada**

Fundação pública vinculada à Secretaria de Assuntos Estratégicos da Presidência da República, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidente

Marcio Pochmann

Diretor de Desenvolvimento Institucional

Fernando Ferreira

Diretor de Estudos e Relações Econômicas e Políticas Internacionais

Mário Lisboa Theodoro

Diretor de Estudos e Políticas do Estado, das Instituições e da Democracia

José Celso Pereira Cardoso Júnior

Diretor de Estudos e Políticas Macroeconômicas

João Sicsú

Diretora de Estudos e Políticas Regionais, Urbanas e Ambientais

Liana Maria da Frota Carleial

Diretor de Estudos e Políticas Setoriais, de Inovação, Regulação e Infraestrutura

Márcio Wohlers de Almeida

Diretor de Estudos e Políticas Sociais

Jorge Abrahão de Castro

Chefe de Gabinete

Fabio de Sá e Silva

Assessor-chefe de Imprensa e Comunicação

Daniel Castro

Corpo Editorial

URL: <http://www.ipea.gov.br>

Editor

Marco Antônio Freitas de Hollanda Cavalcanti

Ouvidoria: <http://www.ipea.gov.br/ouvidoria>

Coeditor

Maurício Cortez Reis

Membros

Alexis Maka

Carlos Henrique Leite Corseuil

Eduardo Pedral Sampaio Fluzza

Eustáquio José Reis

Fernando Veloso

Lauro Roberto Albrecht Ramos

Miguel Nathan Foguel

Pedro Cavalcanti Ferreira

Rodrigo Reis Soares

Secretária-Executiva

Angélica Ferreira de Barros

pesquisa e planejamento econômico - ppe

ipea

Pesquisa e Planejamento Econômico v. 1 – n.1 – jun. 1971.

Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada, 1990 – v. – quadrimestral.

Título anterior: Pesquisa e Planejamento v. 1, n. 1 e 2, 1971

Periodicidade anterior: semestral de 1971–1975.

1. Economia – Pesquisa – Periódicos. 2. Planejamento Econômico – Brasil. I. Brasil.
Instituto de Pesquisa Econômica Aplicada.

ISSN - 0 100-0551

CDD 330.05
33(81) (05)

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou o da Secretaria de Assuntos Estratégicos.

É permitida a reprodução dos textos deste volume e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

SUMÁRIO

**RISCO DAS INSTITUIÇÕES FINANCEIRAS E O
IMPACTO SOBRE A ATIVIDADE ECONÔMICA: UMA
ANÁLISE COM DADOS EM PAINEL PARA O BRASIL 5**
Helder Ferreira de Mendonça, Délio José Cordeiro
Galvão e Renato Falci Villela Loures

**DIREITOS DE PROPRIEDADE E BEM-ESTAR:
AVALIAÇÃO DO IMPACTO DO PROGRAMA
DE REGULARIZAÇÃO FUNDIÁRIA NA QUINTA DO CAJU 29**
Maria Isabel de Toledo Andrade e Valéria Pero

**TRABALHO MATERNO E DESEMPENHO
EDUCACIONAL DAS CRIANÇAS: UMA ANÁLISE
DA PROBABILIDADE DE APROVAÇÃO ESCOLAR 71**
Juliana Maria de Aquino e Elaine Toldo Pazello

**GASTO EM P&D E PODER DE MERCADO:
TEORIA E EVIDÊNCIA PARA O BRASIL 101**
Eduardo Correia e Rodrigo Moita

**UM MODELO ECONOMÉTRICO COM
PARÂMETROS VARIÁVEIS PARA A CARGA
TRIBUTÁRIA BRUTA BRASILEIRA TRIMESTRAL 133**
Mario Jorge Cardoso de Mendonça,
Adolfo Sachsida e Luis Alberto Medrano

RISCO DAS INSTITUIÇÕES FINANCEIRAS E O IMPACTO SOBRE A ATIVIDADE ECONÔMICA: UMA ANÁLISE COM DADOS EM PAINEL PARA O BRASIL*

Helder Ferreira de Mendonça**

Délio José Cordeiro Galvão***

Renato Falci Villela Loures****

Este artigo apresenta evidências empíricas para a presença do *trade-off* entre a atividade econômica e o risco que as instituições financeiras estão dispostas a correr em suas atividades. Por meio da aplicação da metodologia de dados em painel para 66 firmas bancárias brasileiras, foi observado que a adoção de estratégias de maior risco no mercado está associada a maiores índices de lucratividade. Além disso, as evidências empíricas revelam a existência de uma relação positiva entre o hiato do produto e o risco das firmas bancárias.

1 INTRODUÇÃO

Um dos principais objetivos do Novo Acordo de Basileia, implementado em 2004, é o de tentar reduzir os incentivos às instituições financeiras para que assumam posições em risco cada vez mais elevadas, na expectativa de auferir maior lucratividade. Em consequência, surge um delicado embate no setor financeiro e bancário em que ao regulador cabe impedir o surgimento de crises sistêmicas, enquanto o mercado busca meios de alcançar a maior rentabilidade possível assumindo riscos minimamente aceitáveis (ESTRELLA, 2004).

A publicação do Novo Acordo de Basileia trouxe incentivo ao setor bancário para melhorar as práticas de gerenciamento do risco destas instituições (BIS, 2004). Ao tornar o capital regulamentar sensível ao grau de risco assumido pelas instituições financeiras, as medidas propostas no Novo Acordo podem implicar o prolongamento ou a redução da extensão dos ciclos econômicos (CHIANAMEA, 2005). Em outras palavras, os ciclos econômicos podem ser alongados em função dos requerimentos mínimos de capital. Por exemplo, em períodos de recessão os bancos são levados a restringir suas linhas de crédito, pois o risco de perdas por

* Agradecemos aos profícuos comentários realizados por dois pareceristas anônimos. Como de praxe, as possíveis omissões ou imprecisões são de inteira responsabilidade dos autores. As opiniões expressas neste trabalho são exclusivamente dos autores e não refletem a visão do Banco Central do Brasil (BCB).

** Professor Titular do Departamento de Economia da Universidade Federal Fluminense (UFF) e pesquisador do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

*** Analista do BCB e doutorando em Economia da UFF.

**** Doutorando em Economia da UFF e bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

inadimplência é maior. Logo, a restrição de crédito poderá retardar uma possível fase de retomada do crescimento econômico (ALLEN; SAUNDERS, 2004).

Mesmo antes da crise do *subprime* nos Estados Unidos, a ideia de que os requerimentos mínimos de capital pudessem implicar o alongamento dos ciclos econômicos era bastante difundida. Diversos autores conseguiram relacionar a característica pró-cíclica do capital requerido para cobertura dos riscos em operações de crédito às fases de crescimento e recessão da economia (ALLEN; SAUNDERS, 2004; GORDY; HOWELLS, 2004; BORIO, 2004; CHIANAMEA, 2005).

Em análise realizada para o caso brasileiro referente à relação entre a capitalização e a exposição a riscos da indústria bancária, Araújo, Jorge Neto e Linhares (2008) encontraram evidências de que um maior nível de capital amplia a exposição a riscos dos bancos. Tal resultado está em consonância com aqueles encontrados por Koehn e Santomero (1980) e Kim e Santomero (1988). Em particular, permite conjecturar que a solução proposta por Borio (2004), concernente à criação de um colchão de liquidez nas fases de expansão da economia, possa ser utilizada como ferramenta anticíclica capaz de prevenir a formação de novas bolhas no mercado financeiro.

Passada a fase mais aguda da crise do *subprime* é possível identificar, nas principais propostas para revisão das regras de regulação e supervisão bancária, um maior destaque aos efeitos macroeconômicos da regulação financeira.¹ Observa-se uma tendência à substituição do atual modelo de provisionamento de capital para cobertura de perdas das instituições financeiras baseado em perdas pretéritas, por um mecanismo que leve em conta as perdas esperadas ao longo do ciclo econômico. Assim, as provisões para perdas deverão ser calculadas com base na probabilidade de inadimplência inerente ao ciclo econômico em curso e não na probabilidade medida a cada instante.

No Brasil, uma das propostas para limitar a pró-ciclicidade do atual modelo de regulação prudencial é estabelecer um requisito adicional de capital regulamentar, acima do requisito mínimo de 11%, para formar um “colchão” de capital na fase crescente dos ciclos econômicos. Conforme apontado por Tombini (2009), em uma situação que envolva nova crise no sistema financeiro, o requisito adicional poderá ser reduzido ou mesmo eliminado, mitigando os efeitos da crise sobre o restante da economia.

1. Ver: i) FSB (2009); ii) Brunnermeier *et al.* (2009); iii) IMF, (2009); iv) FSA, (2009); v) WEF, (2009); vi) G20 Working group 1, (2009); e vii) Group of Thirty (2009).

O principal objetivo deste artigo é analisar empiricamente, para o caso brasileiro, as relações entre risco e lucratividade das instituições financeiras, bem como entre o nível de produto agregado e o risco na indústria bancária. Para tanto, foram utilizados dados trimestrais para o período compreendido entre o primeiro trimestre de 2001 e o segundo trimestre de 2009 a partir de 66 firmas bancárias. Além desta introdução, o artigo conta com mais quatro seções. A seção 2 oferece uma resenha da literatura ressaltando a relação entre regulação e ciclos econômicos. A seção 3 apresenta os dados e a metodologia utilizados no estudo. A seção 4 mostra as evidências empíricas obtidas a partir da aplicação da metodologia de dados em painel – Método Generalizado dos Momentos – do inglês Generalized Method of Moments (GMM). Por último são apresentadas as conclusões do trabalho.

2 REGULAÇÃO E CICLOS ECONÔMICOS

Com o objetivo de apresentar uma evolução do processo de regulação prudencial, a pró-ciclicidade dos riscos das instituições financeiras e o papel desempenhado pela regulação no Brasil e seus efeitos diante da crise do *subprime*, esta seção realiza uma breve resenha dos principais argumentos que permeiam a literatura sobre o assunto.

2.1 Regulação prudencial

As medidas adotadas nos Estados Unidos, em decorrência da corrida bancária observada depois da quebra da Bolsa de Nova York, em outubro de 1929, podem ser consideradas como as que iniciaram o processo de regulação bancária. Em 1933, o Congresso norte-americano aprovou a segunda Lei Glass Steagall (LGS) que se baseou em duas medidas básicas:² o estabelecimento de uma rígida regulação prudencial, que proibia a constituição dos bancos universais;³ e a formação de uma rede de segurança, que assegurasse os depósitos efetuados nas instituições bancárias e que impedisse ameaças à estabilidade do sistema financeiro (BERNANKE, 2000).

O princípio por trás da proibição à constituição dos bancos universais estava na crença de que, passando a atuar em diferentes áreas do mercado, os bancos ficavam expostos aos riscos inerentes a esses novos mercados. Ademais, a ocorrência de um choque em determinada área do mercado financeiro poderia se propagar e contaminar as instituições de todo o sistema. Assim, a estratégia de especialização dos bancos, além de facilitar a compatibilização em termos de maturidade e liquidez dos recursos captados e os ativos da instituição, dificultaria a propagação de choques.

O outro objetivo que motivou publicação da LGS foi criar uma rede de segurança para o sistema. Destarte, foi criado o Federal Deposit Insurance Corporation (FDIC) com o intuito de assegurar os depósitos nas instituições financeiras. Ademais,

2. A primeira LGS, editada em 1932, não passou de uma tentativa de expansão de crédito.

3. Bancos universais são aqueles autorizados a operar simultaneamente em quaisquer segmentos de mercado.

para prover maior estabilidade ao mercado e impedir uma “corrida bancária” devido a eventos localizados capazes de gerar um risco sistêmico, a presença de um emprestador de última instância tornou-se vital (KRUGMAN, 2009).⁴

O seguro de depósito proporcionado pelo FDIC é considerado o principal sistema de segurança adotado pelos órgãos de supervisão que compõem a chamada regulação de proteção (FLANNERY, 1998). Tem sido a medida mais eficaz para evitar novas corridas bancárias. Embora tenha pouco poder de cobertura diante do reduzido volume de recursos em reserva, a garantia de reposição aos depositantes que detenham valor igual ou inferior a US\$ 250 mil,⁵ em até três dias da decretação de falência do banco, fez com que nem mesmo na fase mais aguda da crise do *subprime* houvesse uma corrida bancária para saques de depósitos à vista.⁶

A estabilidade do sistema bancário norte-americano, a partir da publicação da LGS, foi baseada em três principais medidas: *i*) a segmentação dos mercados financeiros; *ii*) a regulação prudencial; e *iii*) a criação de uma rede de proteção. A rigor, um banco poderia operar sem a formação de capital próprio ao atuar apenas como intermediário entre os agentes de captação e empréstimo.

A partir dos anos 1960, o mercado financeiro foi abalado por diversos choques externos, entre os quais as duas crises do petróleo (1973 e 1979). Em paralelo deu-se o colapso do sistema de Bretton Woods, entre 1971 e 1973, e a maioria dos países industrializados adotou políticas monetárias contracionistas. É importante notar que o resultado da maior volatilidade dos preços, das taxas de câmbio e juros associados a um processo de desregulação financeira levou à liberalização de movimentos de capitais internacionais e ao aumento de incerteza no mercado financeiro (ALLEN; GALE, 2006).

Brasil, México e Argentina foram alguns dos países, da América Latina, mais atingidos pelas crises externas entre os anos 1960 e 1980. Outrossim, o impacto desses fatores externos foi de tal magnitude que colocou em risco os ativos de grandes bancos de Nova York que haviam feito vultosos empréstimos àqueles governos. Temia-se que, no caso de inadimplência, as perdas dos bancos “grandes demais para quebrar”⁷ pudessem implicar um risco sistêmico. Destarte, o governo

4. Para o risco de que dificuldades de bancos individuais possam ter reflexo em todo o sistema bancário e, eventualmente, em toda a economia, ver De Bandt e Hartmann (2000) e Diamond e Dybvig (1983).

5. A partir de 2005 este valor é corrigido quinquenalmente pela taxa acumulada de inflação.

6. No Brasil, o Fundo Garantidor de Crédito (FGC) foi criado e regulamentado pela Resolução do Conselho Monetário Nacional (CMN) nº 2.211 de 1995. Atualmente é regulamentado pela Resolução nº 3.215, de 16 de dezembro de 2004 e seguintes. É gerido por uma sociedade de direito privado e destina-se à cobertura, limitada a R\$ 60 mil, de titulares de depósitos e de determinados investimentos financeiros, em caso de falência, intervenção ou liquidação extrajudicial de instituições financeiras (ALVES; ALVES, 2010).

7. O limite para o tamanho dos bancos nos Estados Unidos foi definido pela Lei Riegle-Neal, em 1994. Segundo essa lei, nenhum banco pode ter mais que 10% dos depósitos de varejo, valor máximo acima do qual poderia causar uma crise sistêmica.

americano interveio junto ao Congresso e obteve os recursos necessários para que o Fundo Monetário Internacional (FMI) pudesse socorrer os países em crise. Em contrapartida foi exigido dos bancos beneficiados pelo empréstimo que constituíssem um valor mínimo de capital próprio do banco. Este procedimento deu origem ao coeficiente de capital composto por capital da própria instituição.

A primeira tentativa supranacional de implantar uma supervisão e um controle na indústria bancária mundial foi feita com a publicação do Acordo de Basileia em 1988 (Acordo de Capital). A implementação dos princípios contidos nesse acordo tinha por objetivo garantir a segurança e a solidez do sistema financeiro mundial a partir da exigência de um capital mínimo regulatório. Esse capital deveria ser suficiente para reduzir o risco de insolvência que pudesse levar a um risco sistêmico, além de suprir eventuais perdas capazes de acarretar prejuízo aos depositantes.

A especialização funcional que vigorava no sistema bancário norte-americano sofria desgastes e, em consequência, ainda nos anos 1980 os bancos universais voltaram a operar. A Europa Ocidental mantinha regimes regulatórios menos exigentes, e a expansão da atividade bancária em um mundo cada vez mais globalizado permitiu maior diversificação das atividades dos bancos. Ainda no mesmo período os bancos passaram a securitizar⁸ seus créditos como forma de diminuir a exposição ao risco de suas carteiras.

Durante mais de dez anos o Acordo de Capital, claramente baseado em uma regulação prudencial de liquidez, foi aceito em âmbito mundial. Entretanto, devido à diversidade de operações realizadas por um único banco, alavancada pelo processo de securitização, as resoluções produzidas pelo Acordo de 1988 mostraram-se ineficientes (HELLMAN; MURDOCK; STIGLITZ, 2000). Conforme observado por Kregel (1998), a classificação de riscos proposta naquele acordo era malfeita, e muito ampla a diferença de riscos dentro de cada categoria. Dado que o custo regulatório era o mesmo, havia o estímulo aos bancos para que procurassem aplicações mais rentáveis, que também eram as que envolviam maior risco.

A entrada do Brasil no grupo de países que aderiram às premissas do Acordo de Capital se deu no ano de 1994 e significou um marco da regulação prudencial no país (ARAÚJO; JORGE NETO; LINHARES, 2008). A partir da publicação da Resolução nº 2.099/1994, passou a ser regra para as instituições financeiras e demais instituições autorizadas funcionar pelo Banco Central do Brasil (BCB) a adoção dos Limites Mínimos de Capital Realizado e Patrimônio Líquido. Ademais, uma série de medidas foi tomada pelos órgãos de regulação e supervisão bancária como forma de respaldar os 25 princípios essenciais de Basileia.⁹

8. Securitizar os créditos é o mesmo que transformá-los em outros produtos financeiros. Esses derivativos podem se espalhar pelo mercado e passar a compor a carteira de ativos de outros bancos como forma de lastrear-lhes novas operações.

9. Ver Core Principles for Effective Banking Supervision. Disponível em: <<http://www.fgvsp.br/institucional/biblioteca/pe/1200300701.pdf>>

O aprimoramento das operações do sistema bancário, a proliferação das operações feitas via internet e a globalização do sistema financeiro levaram o Comitê de Basileia a lançar a proposta do Novo Acordo em 2001. O Novo Acordo (Basileia II), implementado em junho de 2004, estabelece que os bancos indiquem qual parte de seu capital será mantido em disponibilidade para cobertura dos diversos tipos de risco a que a instituição está sujeita. O Novo Acordo baseia-se em três pilares mutuamente complementares: *i*) a definição do coeficiente de capital de risco em função de métodos de avaliação inerentes à estrutura da instituição financeira; *ii*) a definição dos poderes e atribuições do órgão supervisor; e *iii*) o destaque para a disciplina de mercado devido à percepção de que o próprio mercado pode contribuir para manter sua segurança e estabilidade.

A maioria dos países tornou-se signatária ou se comprometeu com os princípios estabelecidos pelo Novo Acordo. Entretanto, os Estados Unidos permaneciam envolvidos com a implementação da Lei Sarbox, que sucedeu aos escândalos corporativos no ano de 2002. Embora essa lei tivesse como principal objetivo aumentar a responsabilidade dos gestores das empresas (incluindo auditorias, advogados etc.), por meio da aplicação de uma série de regras de governança corporativa, ela se mostrou insuficiente para impedir a crise do *subprime*.

2.2 Pró-ciclicidade dos riscos

O atual modelo de regulação financeira possui características pró-cíclicas. As expectativas otimistas formadas durante um ciclo de prosperidade econômica diminuem a aversão ao risco dos investidores. Assim, conforme identificado por Minsky (1982), a maior exposição ao risco torna a economia mais vulnerável a crises financeiras. Com base em um modelo de ciclos de negócios, Gorton (1988) observou que as cinco piores recessões na economia norte-americana foram sucedidas por crises no sistema bancário. Ademais, os resultados encontrados indicam que as crises do sistema financeiro estão relacionadas aos movimentos dos ciclos econômicos que, por sua vez, não ocorrem de forma aleatória.

Allen e Gale (2000) também desenvolveram um modelo que relaciona as crises financeiras aos ciclos de negócios. Segundo as premissas do modelo, a redução no valor dos ativos financeiros dos bancos está relacionada à redução da atividade econômica. Uma das conclusões dos autores é de que a possibilidade de ocorrência de crises serve como freio às ações dos agentes que visem à assunção de maiores riscos para a instituição financeira. De Bandt (1995), tomando como referência a análise desenvolvida por Jacklin e Bhattacharya (1988), elaborou um modelo de informação que utiliza o conceito de incerteza agregada para explicar o risco cíclico dos negócios. Os resultados indicaram que a informação sobre dificuldades

na obtenção de retornos de um único banco pode fazer com que os depositantes iniciem seus saques de forma antecipada dando início a uma corrida bancária.¹⁰

A relação existente entre as crises financeiras e os picos em ciclos de expansão da economia foi analisada por Kindleberger e Aliber (2005). Segundo os autores, as bolhas de ativos se formam e ganham impulso até o instante em que um choque externo interrompa este curso. Neste momento, a realização de lucros, mesmo que pela minoria dos investidores, não é mais compensada com a entrada de novos agentes no mercado. Inicia-se, então, uma corrida para liquidação de posições que acaba culminando com o “estouro da bolha”. Em outras palavras, a quebra de um banco ou de uma grande empresa pode levar ao pânico e à quebra dos mercados financeiros.

Allen e Saunders (2004) mostram que as intermediações financeiras são intensificadas nos períodos de expansão econômica, tornando o ciclo mais longo. Não obstante, também criam um ambiente propício à formação de “bolhas” que, passada a fase expansionista, poderá ser precedida de rápida desaceleração da economia e culminar com o colapso do preço dos ativos e o racionamento de crédito. Assim, em um cenário de expansão da economia, a natureza pró-cíclica da regulação prudencial poderá levar o sistema a uma perda de eficiência. Ademais, durante recessões esta dinâmica implica perda de liquidez para o mercado, o que poderá inviabilizar o retorno do investimento e a retomada do crescimento econômico.

2.3 Regulação e a crise no Brasil

Em 2008, os canais de transmissão da crise do *subprime* para o Brasil foram os mesmos observados na crise da década de 1930.¹¹ Uma crise externa, iniciada no sistema bancário, a partir da quebra do banco americano Lehman Brothers, em setembro de 2008, que se propagou para os demais mercados financeiros por meio de uma abrupta interrupção das linhas de crédito. Entretanto, ao contrário da situação em que se encontrava o país durante a eclosão de outras crises, o Brasil da crise do *subprime* é um país com excesso de regulação do mercado financeiro. Além disso, conforme destacado por Franco (2009), as empresas encontravam-se “subalavancadas” e a crise aportou em meio a um ciclo econômico com claros sinais de superaquecimento.

Um ponto importante que marcou a economia brasileira durante a crise do *subprime* refere-se ao papel exercido pelos bancos públicos. Enquanto os bancos

10. Para uma revisão na literatura sobre corridas bancárias com informação assimétrica, ver Jacklin e Bhattacharya (1988), Chari e Jagannathan (1988), Gorton (1988) e Calomiris e Gorton (1991).

11. O *crash* da Bolsa de Nova York, em 1929, alcançou as economias dos países da América Latina, pelo efeito contágio, a partir do corte nas linhas de crédito norte-americano, que implicaram a paralisação das exportações e a queda dos preços das *commodities* e produtos agrícolas, principais produtos de exportação destas economias. Em consequência, a brusca queda dos preços internacionais do café fez ruir o regime monetário estabelecido no Brasil e deu início a uma fase de grandes desvalorizações da taxa de câmbio (CAMPA, 1990).

privados reduziram de forma drástica a concessão de créditos no início da fase de quebras no mercado financeiro ao redor do mundo, o governo brasileiro adotou política anticíclica ao fazer com que as instituições financeiras oficiais – Banco do Brasil (BB), Caixa Econômica Federal (CAIXA) e Banco Nacional de Desenvolvimento Econômico e Social (BNDES) – ampliassem o volume de empréstimos e liquidez para o mercado. Não obstante, os bancos públicos, assim como os privados, têm respeitado os limites de Basileia (11 vezes o valor patrimonial do banco para utilização em empréstimos) e estão sujeitos ao recolhimento de empréstimos compulsórios tanto sobre depósitos à vista quanto a prazo.

É importante notar que a adoção do regime de metas de inflação (junho de 1999) implicou um incremento na transparência do sistema financeiro e uma condução da política monetária, o que, por conseguinte, contribuiu para a estabilização macroeconômica brasileira.¹² O resultado da maior estabilidade permitiu um rápido desenvolvimento do mercado financeiro e fez com que os agentes privados, na busca de lucros mais elevados, migrassem para investimentos mais arriscados – mercado de ações, débitos subordinados etc. (ver MENDONÇA; LOURES, 2009). Desarte, as transformações observadas no mercado financeiro passaram a demandar o acompanhamento das regras de regulação prudencial. Desde 2001 quase todos os tipos de risco das instituições financeiras apresentam alguma regulamentação própria com capacidade para afetar a avaliação dos riscos nas instituições bancárias (ARAÚJO; JORGE NETO; LINHARES, 2008).¹³

3 DADOS E METODOLOGIA

Para uma análise empírica referente ao risco e à lucratividade das instituições financeiras, bem como do produto e o risco na indústria bancária, foram utilizados dados trimestrais para o período compreendido entre o primeiro trimestre de 2001 e o segundo trimestre de 2009.¹⁴ As informações foram obtidas a partir de 66 firmas bancárias brasileiras, totalizando 2.244 observações para dados em painel.¹⁵ As seguintes variáveis foram empregadas no estudo:

1. Lucro líquido (LL) – representa a divisão entre o lucro líquido (l) e os ativos menos intermediação das instituições financeiras (AI), isto é,

$$LL = \left(\frac{l}{AI} \right) + 1 \quad (1)$$

12. Para uma análise sobre a adoção de metas de inflação e suas consequências no Brasil, ver Mendonça (2007).

13. Para aprofundar-se na formação do arcabouço da regulação prudencial no Brasil, ver Alves e Alves (2010).

14. O início do período de observação dos dados (primeiro trimestre de 2001) está relacionado com a publicação da Resolução nº 2.682/1999 do BCB, que institui a obrigatoriedade do cálculo do índice de Basileia.

15. As instituições financeiras analisadas encontram-se dispostas no quadro A.1 do Apêndice A.

2. Hiato do produto (X) – corresponde à divisão entre o Produto Interno Bruto (PIB)(X^O) (preços de 2000) divulgado pelo Sistema de Contas Nacionais do Instituto Brasileiro de Geografia e Estatística (SCN/IBGE) e o produto potencial (X^P).¹⁶ Assim:

$$X = X^O / X^P \quad (2)$$

3. Índice de Basileia (IB) – constitui uma *proxy* de avaliação do risco para as instituições financeiras. Um índice elevado revela maior solvabilidade e alto nível de capitalização do banco.

O fator mínimo de capital exigido pelo órgão de regulação no Brasil é de 11% das exposições líquidas (o Comitê de Basileia estabelece um mínimo de 8%) e obedece à Resolução nº 2.682/1999, que prescreve percentuais mínimos de provisionamento de acordo com um critério de classificação. À medida que o IB aproxima-se do valor-limite de 11% (estabelecido pelo BCB), o órgão regulador exige a recomposição de capital próprio da firma bancária para que esta possa continuar operando no mercado. Em suma, quanto maior o IB , mais solvente é a instituição bancária, ou seja, menos exposto ao risco e mais capitalizado está o banco.¹⁷

Com base na literatura (ver KAMINSKY; REINHART, 1999), as seguintes variáveis de controle foram utilizadas nos modelos: *i*) taxa de juros *Over-SELIC* (*Selic*); *ii*) Ibovespa (*Ibov*); e *iii*) taxa de câmbio (*Cambio*).¹⁸ A tabela 1 apresenta as estatísticas descritivas das séries.

TABELA 1
Estatísticas descritivas

	<i>IB</i>	<i>Cambio</i>	<i>Ibov</i>	<i>X</i>	<i>LL</i>	<i>Selic</i>
Média	-1,4140	0,8728	10,1571	-0,0001	0,0082	0,0380
Mediana	-1,5950	0,8442	10,1561	0,0024	0,0068	0,0379
Máxima	6,4089	1,2998	11,1193	0,0555	0,1814	0,0564
Mínima	-3,0098	0,5040	9,2021	-0,0690	-0,2842	0,0236
Desvio-padrão	0,7070	0,2005	0,5895	0,0303	0,0245	0,0085
Observações	2.244	2.244	2.244	2.244	2.244	2.244

Fonte: Elaboração dos autores.

16. Devido ao fato de o filtro HP decompor a série de tempo em um componente cíclico e tendência, a tendência obtida pelo filtro HP pode ser interpretada como o produto potencial.

17. O apêndice B apresenta o procedimento de cálculo do índice de Basileia.

18. Foi aplicado logaritmo natural em todas as séries utilizadas nos modelos.

Como forma de eliminar os efeitos não observados nas regressões, este estudo faz a estimativa, em primeira diferença, pelo Método dos Momentos Generalizados – Generalized Method of Moments (GMM) – dinâmico com dados em painel. Conforme ressaltado por Arellano e Bond (1991), uma vantagem deste método em relação aos demais, por exemplo, Mínimos Quadrados Ordinários (MQO) – *Ordinary Least Squares*; e Mínimos Quadrados Generalizados – *Generalized Least Squares* (GLS) –, é que as estimações são confiáveis mesmo no caso de haver variáveis omitidas. Em particular, o uso de variáveis instrumentais permite a estimativa dos parâmetros de forma mais consistente, mesmo no caso de endogeneidade nas variáveis explanatórias e na ocorrência de erros de mensuração (BOND; HOEFFLER; TEMPLE, 2001).

Para analisar as duas relações (lucro líquido e índice de Basileia, e produto e índice de Basileia) são estimados quatro modelos que levam em consideração o tamanho das instituições bancárias, dispostos da seguinte forma:

1. Painel 1 – reúne os nove maiores bancos brasileiros. Juntas, essas instituições representam 90% dos ativos menos o total de intermediação dos bancos considerados neste estudo. Em sua maioria são conglomerados financeiros.
2. Painel 2 – composto por 21 bancos considerados intermediários. Representa cerca de 8% dos ativos menos o total de intermediação dos bancos. Em sua maioria são bancos de investimento.
3. Painel 3 – formado pelos 36 menores bancos brasileiros. Representa 2% dos ativos descontando-se a intermediação total dos bancos. Em sua maioria são bancos financiadores de bens de consumo.
4. Painel 4 – são consideradas todas as instituições no estudo – total de 66 bancos.

Dado o objetivo de observar o comportamento dos bancos na economia brasileira pelo tamanho dos seus ativos, decidiu-se destacar os nove maiores bancos (14% do total da amostra) dos demais. Esta divisão permite maior compreensão do comportamento da grande maioria dos bancos brasileiros que, em geral, são pequenos. Além disso, tal divisão torna possível verificar a existência de alguma similaridade para o caso dos bancos, que de fato possuem um ativo muito elevado, em relação aos que ocupam uma posição intermediária.

Com a intenção de corrigir o problema de heterocedasticidade das estimativas, as matrizes de covariância foram estimadas pelo método de White.¹⁹ Para

19. Foi empregada a matriz de correção de White por período. O fato de as matrizes de covariância terem sido previamente corrigidas pelo método de White tornou desnecessário testar a heterocedasticidade nas séries.

verificar a relevância das variáveis instrumentais do modelo foram aplicados testes de restrições às sobreidentificações (teste de Sargan) como proposto por Arellano (2003). Ademais, conforme proposto por Arellano e Bond (1991), dois testes de primeira ordem (m1) e de segunda ordem (m2) de correlação serial são feitos. Por último, para proceder à estimação das regressões foi necessário testar a existência de estacionariedade das séries.

É importante salientar que mesmo com as premissas do GMM, quando não há correlação na primeira diferença dos regressores endógenos, é necessário testar a presença de raiz unitária na série. Seguindo a metodologia empregada por Bond, Nauges e Windmeijer (2005) foram criadas várias técnicas para testar a raiz unitária para dados em painel. Os testes utilizados neste artigo são: Levin-Lin-Chu (LLC), Im-Pesaran-Shin (IPS), Fisher-ADF (ADF), e Fisher-PP (PP). O teste LLC assume a existência de uma única raiz unitária comum a todos as cortes transversais. Os demais testes assumem a existência de diferentes raízes unitárias nos diferentes cortes transversais. A hipótese nula em todos os testes mencionados considera a não existência de estacionariedade nas séries. Os testes foram realizados para séries em nível, utilizando-se a seleção de defasagens por meio dos critérios de informação de Schwarz. A tabela A.1 (vide apêndice A) apresenta os resultados dos testes de raízes unitárias. Destarte, as séries *LL*, *IB* e *X* foram utilizadas em nível, enquanto as séries *Selic*, *Cambio* e *Ibov* foram consideradas em primeira diferença.

4 EVIDÊNCIAS EMPÍRICAS

Considerando-se a relação entre o *LL* e o *IB* das firmas bancárias foram estimados quatro modelos distintos, de acordo com o tamanho das instituições financeiras, tal como mencionado na seção anterior. A equação (3) é utilizada para os quatro painéis estimados:²⁰

$$LL_t = \beta_0 LL_{t-1} + \beta_1 IB_{t-1} + \beta_2 \Delta Selic_{t-2} + \beta_3 \Delta Ibov_t + \beta_4 X_t + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2) \quad (3)$$

A tabela 2 apresenta os resultados das regressões. Os quatro painéis apresentam estatísticas de Sargan satisfatórias, pois é aceita a hipótese nula de existência de sobreidentificação em todos os modelos, o que valida as variáveis instrumentais utilizadas. Os testes de autocorrelação serial de primeira ordem (m1) e segunda ordem (m2) rejeitam a hipótese de existência de autocorrelação serial nas quatro especificações estimadas.

20. Os modelos foram selecionados com base no princípio da parcimônia. Em relação aos instrumentos, além da defasagem das variáveis explicativas, o agregado monetário M2 (vide KAMINSKY; REINHART, 1999) também foi utilizado.

TABELA 2

Estimação: painel dinâmico – efeito sobre o lucro líquido das firmas bancárias

	Painel 1		Painel 2		Painel 3		Painel 4	
	Coeficiente	Estatística-t	Coeficiente	Estatística-t	Coeficiente	Estatística-t	Coeficiente	Estatística-t
LL_{t-1}	0,154	4,464***	0,272	8,776***	0,255	157,125***	0,230	327,470***
IB_{t-1}	-0,015	-4,547***	-0,003	-2,231**	-0,004	-8,527***	-0,001	-5,225***
$\Delta Selic_{t-2}$	2,749	1,665*	0,288	1,516	0,029	1,702*	0,081	11,864***
$\Delta lbov_t$	-0,021	-1,059	0,007	3,935***	-0,017	-28,364***	-0,011	-83,868***
X_t	0,069	4,314***	0,061	16,637***	0,070	28,941***	0,065	96,051***
Número de instrumentos	16		13		12		13	
Número de observações	252		588		1.008		1.848	
Teste de Sargan	4,868		17,250		33,095		65,449	
(valor-p)	0,301		0,370		0,365		0,325	
m1	-3,261		-9,117		-5,822		-5,962	
(valor-p)	0,001		0,000		0,000		0,000	
m2	-0,206		1,551		-0,215		-0,163	
(valor-p)	0,837		0,122		0,829		0,870	

Fonte: Elaboração dos autores.

Nota: *** significativo a 1%; ** significativo a 5%; e * significativo a 10%.

O coeficiente referente ao índice de Basileia é negativo e possui significância estatística nos quatro modelos estimados. Logo, não há dúvida de que, independentemente do tamanho das instituições financeiras consideradas, uma solvência maior dos bancos está associada a uma menor lucratividade. Este resultado é mais forte para o caso dos grandes bancos (painel 1), o que sugere que a busca por lucros mais elevados tornam essas instituições mais arriscadas. Essa é uma observação importante, pois o fato de serem consideradas “grandes demais para quebrarem” pode fragilizar o sistema financeiro criando a possibilidade de uma crise sistêmica.

Os resultados indicam uma relação positiva entre o lucro líquido dos bancos no período corrente com o do período passado (LL_{t-1}). Em outras palavras, a obtenção de elevada lucratividade em um período tende a alavancar resultados positivos nos períodos subsequentes.

Observa-se que, em todos os painéis, a taxa de juros básica possui relação positiva com o lucro líquido das firmas bancárias. Esse resultado revela que elevações na taxa de juros alavancam maiores lucros para o setor financeiro. O coeficiente da $Selic_{t-2}$ no primeiro painel sugere a existência de uma elevada participação de títulos da dívida pública nos portfólios dos grandes bancos. No segundo painel, a $Selic_{t-2}$ apresentou significância estatística para um nível de confiança de 15% e

um coeficiente abaixo daquele observado para os grandes bancos. Em suma, esse resultado indica um menor volume de títulos da dívida pública no portfólio dos bancos de porte intermediário. Em relação ao painel 3, a relação positiva entre a $Selic_{t-2}$ e o LL é explicada pelo fato de os bancos de menor porte serem em sua maioria financiadores de bens de consumo e pelo fato de esses financiamentos estarem atrelados, em sua maioria, à taxa $Selic$. Logo, elevações na taxa básica de juros tendem a aumentar a lucratividade dos bancos de menor porte.

Por último, observa-se que a variável X_t apresentou uma relação positiva com o lucro líquido das firmas bancárias em todos os modelos. Esse resultado indica que um aquecimento econômico estimula uma lucratividade maior das instituições financeiras.

Os painéis 1, 3 e 4 mostram uma relação negativa entre o Ibovespa e o lucro líquido das instituições financeiras. Não obstante, a relação não possui significância estatística no painel 1. Quanto ao painel 2, o fato de o $Ibov$ apresentar uma relação positiva com o LL sugere que os bancos intermediários possuem em seus portfólios participação elevada em ativos da bolsa de valores, indicando um comportamento mais agressivo dos mesmos. Destarte, ao contrário dos bancos de maior porte, que possuem grande volume de títulos da dívida pública atrelados à taxa do Sistema Especial de Liquidação e de Custódia (SELIC), os bancos intermediários possuem maior volume de ações em suas carteiras, o que, por conseguinte, está atrelado a um maior risco. No painel 3 a relação negativa entre o $Ibov$ e o LL é explicada pelo fato de os bancos na amostra serem, na maioria, bancos financiadores de bens de consumo. Em suma, uma elevação do $Ibov$ pode gerar um efeito riqueza que estimula os indivíduos a reduzirem financiamentos por meio do aumento de compras à vista. Logo, o resultado desse processo é uma queda no lucro dos bancos de menor porte.

Da mesma forma que no procedimento adotado para o caso anterior, a análise da relação entre o hiato do produto (X) e o índice da Basileia (IB) é feita por meio de quatro painéis levando em conta o tamanho das instituições bancárias. A equação utilizada nas estimações é dada por:²¹

$$X_t = \alpha_0 X_{t-1} + \alpha_1 IB_{t-1} + \alpha_2 \Delta Selic_{t-2} + \alpha_3 \Delta Cambio_{t-3} + v_t \quad v_t \sim N(0, \sigma^2) \quad (4)$$

Os testes de correlação serial de primeira e segunda ordem para os modelos presentes na tabela 3 indicam a inexistência do problema de autocorrelação. Além disso, o resultado dos testes de Sargan aponta para a validade dos instrumentos utilizados.

21. As observações presentes na nota de rodapé 20 também são válidas para este caso.

TABELA 3

Estimação: painel dinâmico – efeito sobre o hiato do produto

	Painel 1		Painel 2		Painel 3		Painel 4	
	Coeficiente	Estatística-t	Coeficiente	Estatística-t	Coeficiente	Estatística-t	Coeficiente-net	Estatística-t
X_{t-1}	-0,100	-5,573***	-0,111	-12,928***	-0,130	-52,129***	-0,017	-17,029***
IB_{t-1}	-0,111	-11,444***	-0,097	-17,483***	-0,048	-20,640***	-0,034	-40,179***
$\Delta Selic_{t-2}$	-2,702	-19,010***	-2,620	-43,093***	-2,503	-100,957***	-2,503	-241,603***
$\Delta Cambio_{t-3}$	0,055	6,520***	0,046	18,993***	0,045	31,354***	0,129	133,204***
Número de instrumentos	15		14		14		43	
Número de observações	252		609		1.044		1.452	
Teste de Sargan	8,929		20,950		35,977		65,998	
(valor-p)	0,112		0,229		0,143		0,108	
m1	-4,438		-3,862		-3,681		-3,224	
(valor-p)	0,000		0,000		0,000		0,001	
m2	1,519		0,633		0,327		0,281	
(valor-p)	0,130		0,527		0,744		0,779	

Fonte: Elaboração dos autores.

Nota: *** significativo a 1%; ** significativo a 5%; e * significativo a 10%.

Observa-se que, em todas as especificações, os coeficientes do índice de Basileia apresentam uma relação negativa com o hiato do produto e possuem significância estatística. Em outras palavras, uma menor exigência de capital das instituições financeiras contribui para uma ampliação da atividade econômica. O coeficiente é mais elevado para o caso dos grandes bancos e, portanto, indica que uma posição conservadora dos mesmos pode provocar uma desaceleração econômica. Em suma, o resultado encontrado é muito importante, pois revela a existência de um *trade-off* entre produto e risco financeiro.

Os resultados presentes na tabela 3 mostram a existência de uma relação negativa entre o hiato do produto corrente (X_t) e o hiato do produto do período anterior (X_{t-1}). A provável explicação para esse resultado é a de que o produto tende para o potencial; logo, um aumento no hiato do produto em determinado período tende a se reduzir no período subsequente.

Todos os painéis denotam relação negativa entre a $Selic_{t-2}$ e o hiato do produto. Este resultado está em consonância com a teoria dos mecanismos de transmissão monetária, no qual elevações na taxa básica de juros promovem quedas no produto. A validade da teoria supracitada também é atestada para a taxa de câmbio. A relação positiva e com significância estatística para os quatro painéis indica que

desvalorizações cambiais levam a um melhor desempenho das transações correntes, o que, por sua vez, acarreta aumento do produto.

5 CONSIDERAÇÕES FINAIS

Uma novidade apresentada neste artigo é a inserção da regulação prudencial, capturada pelo índice de Basileia, para observar os efeitos de uma exposição maior ao risco dos bancos sobre a atividade econômica. Em particular é observado que, na busca por maiores lucros, os bancos ficam sujeitos a uma exposição maior ao risco. Destarte, devido ao menor rigor na concessão de crédito, o volume de crédito disponível tende a aumentar no mercado. Em consequência, há um aumento na liquidez da economia que favorece novos investimentos e, portanto, é gerado um estímulo para o crescimento econômico.

As evidências empíricas obtidas sugerem que instituições que adotam estratégias de atuação no mercado envolvendo maior risco obtêm maiores índices de lucratividade. Além disso, a comprovação da existência de uma relação positiva entre o hiato do produto e o risco das firmas bancárias no Brasil indica a existência de um *trade-off* entre o risco bancário e o produto. Em outras palavras, os resultados encontrados estão alinhados com os trabalhos que identificam uma relação negativa entre os níveis de capitalização e de risco.

Segundo as conclusões encontradas em Araújo, Jorge Neto e Linhares (2008), o resultado supracitado está em consonância com a ideia de que o capital regulatório pode ser uma ferramenta capaz de preparar o sistema financeiro para enfrentar futuras crises. Não obstante, é de se esperar que a redução da alíquota de requerimento de capital nos períodos de crise nos sistemas financeiros constitua um mecanismo que amorteça os ciclos. Em suma, os resultados apresentados neste artigo sinalizam para a importância do índice de Basileia como ferramenta de regulação prudencial capaz de equilibrar os ciclos econômicos e mitigar a exposição ao risco da indústria bancária em mercados como o brasileiro.

ABSTRACT

This article shows empirical evidence for the presence of the trade-off between economic activity and the risk that financial institutions are willing to incur in their activities. Through the use of panel data methodology for 66 Brazilian banking firms, it was observed that the adoption of riskier strategies in the market is associated with greater profitability. Furthermore, the empirical evidence denotes the existence of a positive relation between the output gap and the risk of banking firms.

REFERÊNCIAS

- ALLEN, L.; SAUNDERS, A. Incorporating systemic influences into risk measurements: a survey of the literature. **Journal of Finance Research**, v. 26, n. 2, p. 161-191, 2004.
- _____; GALE, D. Financial contagion. **Journal of Political Economy**, v. 08, n. 1, p. 1-33, 2000.
- _____; _____. Systemic risk and regulation. In: **NBER CONFERENCE ON THE RISKS OF FINANCIAL INSTITUTIONS**. 2006. Disponível em: <<http://fic.wharton.upenn.edu/fic/papers/05/0524.pdf>>
- ALVES, S. D. S.; ALVES, T. M. S. A experiência brasileira de regulação: um caso de sucesso? In: GIAMBIAGI, F. **Risco e regulação: por que o Brasil enfrentou bem a crise financeira e como ela afetou a economia mundial**. Rio de Janeiro: Campus, 2010. cap. 12, p. 171-188.
- ARAÚJO, L. A. D.; JORGE NETO, P. M.; LINHARES, F. Capital, risco e regulação de bancos no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro: Ipea, v. 38, n. 3, p. 459-486, 2008.
- ARELLANO, M. **Panel data econometrics**. Oxford: Oxford University Press, 2003.
- _____; BOND, S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. **Review of Economic Studies**, v. 58, n. 2, p. 277-297, 1991.
- BERNANKE, B. **Essays on the Great Depression**. Estados Unidos: Princeton University Press, 2000.
- BIS. Bank for International Settlements. **Basel II: international convergence of capital measurement and capital standards**, 2004.
- BOND, S. R.; HOEFFLER, A.; TEMPLE, J. **GMM estimation of empirical growth models**. London: CEPR, 2001 (Discussion Paper, n. 3.048).
- _____; NAUGES, C.; WINDMEIJER, F. **Unit roots: identification and testing in micro panels**. CEMMAP, 2005 (Working Paper, n. CWP07/05).
- BORIO, C. **Procyclicality in Basel II: can we treat the disease without killing the patient?** (Comments on M. Gordy and B. Howells). Presented in Joint Workshop on Accounting, Transparency and Bank Stability, Basel, May 2004.
- BRUNNERMEIER, M. *et al.* The fundamental principles of financial regulation. **Geneva reports on the world economy**. 2009. Disponível em: <<http://www.voxeu.org/reports/Geneva11.pdf>>
- CALOMIRIS, C.; GORTON, G. The origins of banking panics: models, facts and bank regulation. In: HUBBARD, G. (Ed.). **Financial markets and financial crises**. Chicago: University of Chicago Press, 1991.
- CAMPA, J. M. Exchange rates and economic recovery in 1930s: an extension to Latin America. **Journal of Economic History**, v. 50, n. 3, p. 697-682, 1990.
- CHARI, V. V.; JAGANNATHAN, R. Banking panics, information, and rational expectations equilibrium. **Journal of Finance**, v. 43, n. 3, p. 749-761, 1988.
- CHIANAMEA, D. R. Basiléia II e os ciclos econômicos. **Economia Política Internacional: Análise Estratégica**, n. 6, p. 5-14, 2005.
- CMN. **Resolução nº 2.099, de 17/08/1994**. Banco Central do Brasil, 1994.
- _____. **Resolução nº 2.837, de 30/05/2001**. Banco Central do Brasil, 2001.
- DE BANDT, O. Competition among financial intermediaries and the risk of contagious failures. **Documents de Travail 30**, Banque de France, 1995.

_____.; HARTMANN, P. **Systemic risk**: a survey. European Central Bank, 2000 (Working Paper, n. 35).

DIAMOND, D.; DYBVIK, P. Bank runs, deposit insurance and liquidity. **Journal of Political Economy**, v. 91, n. 3, p. 401-419, 1983.

ESTRELLA, A. Bank capital and risk: is voluntary disclosure enough? **Journal of Financial Services Research**, v. 26, n. 2, p. 145-160, 2004.

FLANNERY, M. J. Using market information in prudential bank supervision: a review of the U.S. empirical evidence. **Journal of Money, Credit and Banking**, v. 30, n. 3, p. 273-305, 1998.

FRANCO, G. H. B. O desembarque da crise no Brasil: 10 idéias. In: BACHA, E. L.; GOLDFAJN, I. (Org.). **Como reagir a crise? Políticas econômicas para o Brasil**. Rio de Janeiro: Imago, 2009.

FSA. **The turner review**: a regulatory response to the global banking crisis 2009. Disponível em: <http://www.fsa.gov.uk/pubs/other/turner_review.pdf>

FSB. Report of the financial stability forum on enhancing market and institutional resilience: 2009. Disponível em: <http://www.financialstabilityboard.org/publications/r_0804.pdf>

G20 Working group 1: 2009. **Enhancing sound regulation and strengthening transparency**. Disponível em: <http://www.g20.org/about_working_groups.aspx>

GORDY, M. B.; HOWELLS, B. **Procyclicality in Basel II**: can we treat the disease without killing the patient? Washington, D.C.: Board of Governors of the Federal Reserve System, 2004.

GORTON, G. Banking panics and business cycles. **Oxford Economic Papers**, v. 40, n. 4, p. 751-781, 1988.

GROUP OF THIRTY. **Financial reform**: a reform: a framework for financial stability. 2009. Disponível em: <<http://www.group30.org/pubs/recommendations.pdf>>

HELLMAN, T. F.; MURDOCK, K. C.; STIGLITZ, J. E. Liberalization, moral hazard in banking, and prudential regulation: are capital requirements enough? **American Economic Review**, v. 90, n. 1, p. 147-165, 2000.

IMF **Lessons of the financial crisis for future**. 2009. Disponível em: <<http://www.imf.org/external/np/pp/eng/2009/020409.pdf>>

JACKLIN, C. J.; BHATTACHARYA, S. Distinguishing panics and information based bank runs: welfare and policy implications. **Journal of Political Economy**, v. 96, n. 3, p. 568-592, 1988.

KAMINSKY, G.; REINHART, C. M. The twin crises: the causes of banking and balance-of-payments problems. **American Economic Review**, v. 89, n. 3, p. 473-500, 1999.

KIM, D.; SANTOMERO, M. Risk in banking and capital regulation. **Journal of Finance**, v. 43, n. 5, p. 1.219-1.233, 1988.

KINDLEBERGER C. P.; ALIBER, R. **Manias, panics, and crashes**: a history of financial crisis. Nova York: John Wiley & Sons, 2005.

KOEHN, M.; SANTOMERO, M. Regulation of bank capital and portfolio risk. **Journal of Finance**, v. 35, n. 5, p. 1.235-1.244, 1980.

KREGEL, J. The past and the future of banks. **Quaderni di Ricerche**, Bancaria Editrice, n. 21, 1998.

KRUGMAN, P. **A crise de 2008 e a economia da depressão**. Rio de Janeiro: Campos, 2009.

MENDONÇA, H. F. de. Towards credibility from inflation targeting: the Brazilian experience. **Applied Economics**, v. 39, n. 20, p. 2.599-2.615, 2007.

_____.; LOURES, R. F. V. Market discipline in the Brazilian banking industry: an analysis for the subordinated debt holders. **Journal of Regulatory Economics**, v. 36, n. 3, p. 286-307, 2009.

MINSKY, H. P. **Can it happen again?** Armonk: M. E. Sharpe, 1982.

TOMBINI, A. Lições da crise financeira e reforma regulatória. **Valor Econômico**, 22 out. 2009. Disponível em: <<http://www.jusbrasil.com.br/noticias/1974473/licoes-da-crise-financeira-e-reforma-regulatoria>>

WEF. **The future of the global financial system**. 2009. Disponível em: <<http://www.weforum.org/en/initiatives/Scenarios/NewFinancialArchitecture/index.htm>>

(Originais submetidos em julho de 2009. Última versão recebida em fevereiro de 2011. Aprovada em abril de 2011.)

APÊNDICE A

QUADRO A.1

Relação de instituições bancárias analisadas e classificação por painel

Painel 1	Painel 2		Painel 3			
BB	BNP Paribas	Basa	Bansicredi	Bonsucesso	Cargill	Bancnacion
Itaú	Banrisul	Daycoval	Clássico	Triângulo	Ribeirão Preto	Pottencial
Bradesco	Pactual	Mercantil do Brasil	Barclays Galícia	Fator	Gerdau	La República
CAIXA	BNB	ABC-Brasil	ING	Modal	Credibel	
Santander	Alfa	Sofisa	Schahin	SMBC	Luso Brasileiro	
HSBC	BBM	Rabobank	Indssval	Prosper	Cédula	
Votorantim	Deutsh	Pine	Rural	VR	Renner	
Safra	BIC	Ibibank	Banese	Rendimento	Opportunity	
Citibank	Fibra	Bancoob	GE Capital	Guanabara	Bcomuruguai	
	BMG	Societé Generale	Banpará	Matone	La Provincia	
	Banestes		Industrial do Brasil	Intercap	FICSA	

Fonte: Elaboração dos autores.

TABELA A.1
Testes de raiz unitária

	C ¹						C, T ²						Nenhum ³		
	LLC	IPS	ADF	PP	LLC	IPS	ADF	PP	LLC	IPS	ADF	PP	LLC	ADF	PP
X	Estadística	6,038	-2,670	29,763	277,113	8,779	-0,750	14,167	244,325	-7,384	67,874	437,431			
	Probabilidade	1,000	0,004	0,040	0,000	1,000	0,227	0,718	0,000	0,000	0,000	0,000			
BI	Estadística	-5,271	-5,486	67,904	65,696	-3,740	-3,960	48,127	44,548	-2,306	23,112	24,145			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,011	0,186	0,150			
LL	Estadística	-11,793	-7,833	81,394	199,160	-16,761	-10,115	316,258	216,713	-3,010	57,849	105,878			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000			
Selic	Estadística	-0,125	-0,656	15,919	7,498	-4,547	-5,948	65,766	21,245	-2,479	18,216	15,419			
	Probabilidade	0,450	0,256	0,598	0,985	0,000	0,000	0,000	0,267	0,007	0,442	0,633			
D(Selic)	Estadística	-9,289	-8,188	97,456	32,749	-8,296	-7,255	78,766	18,844	-12,177	148,108	74,175			
	Probabilidade	0,000	0,000	0,000	0,018	0,000	0,000	0,000	0,402	0,000	0,000	0,000			
lbov	Estadística	1,386	3,737	2,022	2,566	-1,174	-3,313	38,125	14,613	3,994	0,876	1,234			
	Probabilidade	0,917	1,000	1,000	1,000	0,120	0,001	0,004	0,688	1,000	1,000	1,000			
D(lbov)	Estadística	-10,123	-7,659	88,877	74,256	-9,758	-7,012	78,039	47,207	-10,656	129,717	123,135			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000			
Cambio	Estadística	0,302	0,587	9,392	11,683	-2,954	-3,147	36,566	29,368	-0,734	9,507	9,653			
	Probabilidade	0,619	0,722	0,950	0,863	0,002	0,001	0,006	0,044	0,232	0,947	0,943			
D(Cambio)	Estadística	-13,100	-11,732	149,145	95,115	-11,522	-10,121	116,719	84,361	-11,842	150,718	141,746			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000			
X	Estadística	9,223	-4,078	69,446	646,597	13,411	-1,145	33,055	570,093	-11,279	158,372	1,020,670			
	Probabilidade	1,000	0,000	0,005	0,000	1,000	0,126	0,837	0,000	0,000	0,000	0,000			
BI	Estadística	-7,510	-8,037	147,367	137,536	-6,108	-5,574	109,917	179,699	-0,655	26,537	26,004			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,256	0,970	0,975			
LL	Estadística	-32,991	-16,148	195,953	337,970	-9,268	-9,913	191,113	622,750	-11,587	394,951	219,843			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000			

(continua)

	C, T ²												Nenthum ³		
	C ¹				C ²				T ²				LLC	ADF	PP
	LLC	IPS	ADF	PP	LLC	IPS	ADF	PP	LLC	IPS	ADF	PP	LLC	ADF	PP
<i>Se/icc</i>	Estadística	-0,191	-1,002	37,143	17,496	-6,946	-9,085	153,454	49,573	-3,787	42,505	35,978			
	Probabilidade	0,424	0,158	0,684	1,000	0,000	0,000	0,000	0,197	0,000	0,449	0,732			
	Estadística	-14,190	-12,508	227,397	76,415	-12,673	-11,083	183,787	43,970	-18,600	345,586	173,074			
	Probabilidade	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,388	0,000	0,000	0,000			
<i>D(Se/icc)</i>	Estadística	2,117	5,708	4,719	5,988	-1,793	-5,061	88,958	34,096	6,101	2,045	2,880			
	Probabilidade	0,983	1,000	1,000	1,000	0,037	0,000	0,000	0,802	1,000	1,000	1,000			
	Estadística	-15,463	-11,700	207,380	173,264	-14,906	-10,710	182,090	110,149	-16,278	302,674	287,314			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000			
<i>D(ibov)</i>	Estadística	0,461	0,897	21,915	27,260	-4,512	-4,807	85,321	68,524	-1,121	22,183	22,524			
	Probabilidade	0,678	0,815	0,996	0,962	0,000	0,000	0,000	0,006	0,131	0,995	0,994			
	Estadística	-20,011	-17,922	348,006	221,936	-17,600	-15,460	272,344	196,842	-18,089	351,675	330,740			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000			
<i>X</i>	Estadística	12,076	-5,339	119,051	1.108,450	17,559	-1,499	56,666	977,302	-14,767	271,496	1.749,720			
	Probabilidade	1,000	0,000	0,000	0,000	1,000	0,067	0,907	0,000	0,000	0,000	0,000			
	Estadística	-3,685	-5,167	147,463	133,522	-3,960	-3,756	130,285	115,067	-0,543	82,973	80,054			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,294	0,177	0,241			
<i>LL</i>	Estadística	-11,466	-12,017	320,196	538,920	-11,813	-10,183	283,841	539,961	-10,161	268,109	414,454			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000			
	Estadística	-0,251	-1,312	63,674	29,993	-9,094	-11,895	263,064	84,982	-4,958	72,866	61,677			
	Probabilidade	0,401	0,095	0,747	1,000	0,000	0,000	0,000	0,141	0,000	0,449	0,802			
<i>Se/icc</i>	Estadística	-18,579	-16,377	389,824	130,997	-16,593	-14,511	315,063	75,378	-24,353	592,432	296,699			
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,370	0,000	0,000	0,000			
	Estadística	2,771	7,473	8,089	10,266	-2,347	-6,627	152,499	58,451	7,987	3,505	4,936			
	Probabilidade	0,997	1,000	1,000	1,000	0,010	0,000	0,000	0,875	1,000	1,000	1,000			

(continua)

(continuação)		C ₁ , T ₂												Nenhum ³		
		C ₁				C ₂				T ₂				LLC	ADF	PP
		LLC	IPS	ADF	PP	LLC	IPS	ADF	PP	LLC	IPS	ADF	PP	LLC	ADF	PP
<i>D(lbov)</i>	Estatística	-20,245	-15,319	355,508	297,025	-19,517	-14,023	312,155	188,826	-21,313	-14,023	312,155	188,826	-21,313	518,869	492,539
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Schwarz ⁴ Painel 3	Estatística	0,603	1,175	37,569	46,731	-5,908	-6,293	146,264	117,470	-1,467	-6,293	146,264	117,470	-1,467	38,028	38,613
	Probabilidade	0,727	0,880	1,000	0,991	0,000	0,000	0,000	0,001	0,071	0,000	0,000	0,001	0,071	1,000	1,000
<i>D(Cambio)</i>	Estatística	-26,201	-23,465	596,582	380,461	-23,044	-20,241	466,876	337,443	-23,684	-20,241	466,876	337,443	-23,684	602,872	566,983
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
<i>X</i>	Estatística	16,351	-7,229	218,260	2,032,160	23,775	-2,030	103,888	1,791,720	-19,995	-2,030	103,888	1,791,720	-19,995	497,742	3,207,830
	Probabilidade	1,000	0,000	0,000	0,000	1,000	0,021	0,966	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
<i>BI</i>	Estatística	-8,391	-9,838	361,403	335,732	-7,173	-6,987	284,648	336,276	-1,980	-6,987	284,648	336,276	-1,980	133,468	130,186
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,024	0,000	0,000	0,000	0,024	0,448	0,528
<i>LL</i>	Estatística	-35,672	-20,875	597,223	1,075,010	-21,697	-16,854	791,125	1,378,050	-15,472	-16,854	791,125	1,378,050	-15,472	720,756	739,740
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
<i>Selic</i>	Estatística	-0,339	-1,777	116,736	54,988	-12,314	-16,106	482,284	155,799	-6,713	-16,106	482,284	155,799	-6,713	133,587	113,074
	Probabilidade	0,367	0,038	0,826	1,000	0,000	0,000	0,000	0,077	0,000	0,000	0,000	0,077	0,000	0,445	0,882
Schwarz ⁴ Painel 4	Estatística	-25,156	-22,174	714,677	240,162	-22,467	-19,648	577,615	138,192	-32,974	-19,648	577,615	138,192	-32,974	1,086,130	543,947
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,339	0,000	0,000	0,000	0,339	0,000	0,000	0,000
<i>lbov</i>	Estatística	3,752	10,119	14,830	18,820	-3,179	-8,972	279,582	107,160	10,815	-8,972	279,582	107,160	10,815	6,426	9,050
	Probabilidade	1,000	1,000	1,000	1,000	0,001	0,000	0,000	0,945	1,000	0,000	0,000	0,945	1,000	1,000	1,000
<i>D(lbov)</i>	Estatística	-27,412	-20,741	651,765	544,545	-26,426	-18,987	572,283	346,181	-28,857	-18,987	572,283	346,181	-28,857	951,261	902,987
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
<i>Cambio</i>	Estatística	0,817	1,591	68,877	85,673	-7,999	-8,521	268,150	215,363	-1,987	-8,521	268,150	215,363	-1,987	69,718	70,791
	Probabilidade	0,793	0,944	1,000	0,999	0,000	0,000	0,000	0,000	0,024	0,000	0,000	0,000	0,024	1,000	1,000
<i>D(Cambio)</i>	Estatística	-35,476	-31,771	1,093,730	697,512	-31,202	-27,407	855,939	618,645	-32,069	-27,407	855,939	618,645	-32,069	1,105,270	1,039,470
	Probabilidade	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Fonte: Elaboração dos autores.

Notas: ¹ Modelo com constante.

² Modelo com constante e tendência.

³ Modelo sem constante e tendência.

⁴ Defasagem de acordo com os critérios de Schwarz: teste LLC – raiz comum – $H_0: \alpha = 0$; teste IPS – raiz individual – $H_0: \alpha = 0$ (para cada i); teste PP – raiz individual – $H_0: \alpha = 0$ (para cada i); teste

APÊNDICE B

CÁLCULO DO ÍNDICE DE BASILEIA

O indicador do IB é calculado a partir da seguinte equação:

$$IB = \frac{PR}{PLE} / F \quad (B.1)$$

O *PR* representa o patrimônio de referência da firma bancária e é definido pelo somatório de dois níveis (CMN, 2001):

1. nível I: representado pelo patrimônio líquido, acrescido do saldo das contas de resultado credoras, e deduzido do saldo das contas de resultado devedoras, excluídas as reservas de reavaliação, as reservas para contingências e as reservas especiais de lucros relativas a dividendos obrigatórios não distribuídos e deduzidos os valores referentes a ações preferenciais cumulativas e a ações preferenciais resgatáveis; e
2. nível II: representado pelas reservas de reavaliação, reservas para contingências, reservas especiais de lucros relativas a dividendos obrigatórios não distribuídos, ações preferenciais cumulativas, ações preferenciais resgatáveis, dívidas subordinadas e instrumentos híbridos de capital e dívida.

O *PLE* representa o patrimônio líquido exigido das firmas em função do risco incorrido em suas atividades operacionais (CMN, 1994). Depois das alterações da Resolução nº 2.099 da CMN o *PLE* passou a ser calculado conforme a equação

$$PLE = F(Apr) + (SWAP) + (PIR) + (EX) \quad (B.2)$$

onde:

Apr – Ativo ponderado pelo risco = total do produto dos títulos do Ativo Circulante e Realizável a Longo Prazo (código 1.0.0.00.00-7 do Plano Contábil das Instituições do Sistema Financeiro Nacional (COSIF) pelos fatores de risco correspondentes + produto do Ativo Permanente (código 2.0.0.00.00-4 do COSIF) pelo fator de risco correspondente + produto dos títulos de Coobrigações e Riscos em Garantias Prestadas (código 3.0.1.00.00-4 do COSIF) pelos fatores de risco correspondentes (CMN, 1994).

SWAP – representa o capital exigido para a cobertura de riscos inerentes a *swap*.

PIR – representa o capital exigido para a cobertura de riscos inerentes à taxa de juros prefixada.

EX – representa o capital exigido para a cobertura de riscos inerentes a variações na taxa de câmbio.

F – representa o fator mínimo do *BI* exigido pelos órgãos reguladores (no Brasil 11% e pelo comitê da Basileia 8%).

DIREITOS DE PROPRIEDADE E BEM-ESTAR: AVALIAÇÃO DO IMPACTO DO PROGRAMA DE REGULARIZAÇÃO FUNDIÁRIA NA QUINTA DO CAJU*

Maria Isabel de Toledo Andrade**

Valéria Pero***

Este trabalho analisa a relação entre direitos de propriedade e bem-estar. Considerando-se a relevância das instituições na economia, supõe-se que, se os direitos de propriedade dos imóveis em comunidades de baixa renda forem mais bem definidos, atribuídos e garantidos, menores serão os custos de transação e, conseqüentemente, maior será o valor do ativo e da renda dos proprietários destes ativos. Além disso, com os direitos de propriedade definidos e assegurados, os moradores de comunidades de baixa renda são incluídos na cidade formal e passam a ter mais acesso a diversos mercados. O presente estudo oferece uma avaliação de impacto do Programa de Regularização Fundiária implementado pela Prefeitura Municipal do Rio de Janeiro na comunidade da Quinta do Caju, que concedeu título de propriedade, registrado em cartório, aos domicílios beneficiados pelo programa. A partir do painel construído com informações socioeconômicas dos mesmos domicílios da comunidade da Quinta do Caju antes e depois do programa, observam-se os efeitos do programa no bem-estar dos moradores. Em particular, por meio do método de *propensity score matching*, constata-se que o programa tem um impacto positivo e estatisticamente significativo na renda domiciliar *per capita* dos domicílios beneficiados pelo programa.

1 INTRODUÇÃO

As favelas são consideradas, em geral, espaços urbanos de ocupação ilegal desprovidos de serviços públicos essenciais. A maioria dos domicílios nas favelas não apresenta os direitos de propriedade bem definidos, atribuídos e assegurados. Diante desse problema, uma das medidas discutidas pela academia e pelos formuladores de políticas públicas, para combater a pobreza e a desigualdade social, é a regularização das ocupações irregulares. Tanto que, no início de 2004, o governo federal, por meio do Ministério das Cidades, apresentou o Programa de Regularização Fundiária com o objetivo de regularizar as ocupações ilegais nas cidades brasileiras.

Em particular na cidade do Rio de Janeiro, a prefeitura, por intermédio da Secretaria Municipal de Habitação, implantou o Programa de Regularização Fundiária, visando a regularização urbanística e fundiária das áreas atingidas pelo

* Este trabalho contou com o financiamento da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) na elaboração da pesquisa de campo na Quinta do Caju, em 2008.

** Doutora em Economia pelo Instituto de Economia (IE) da Universidade Federal do Rio de Janeiro (UFRJ) e pesquisadora em Propriedade Industrial do Instituto Nacional de Propriedade Industrial (Inpi).

*** Professora adjunta do IE/UFRJ.

Programa Favela-Bairro, concluindo assim a última etapa do programa.¹ A Quinta do Caju, comunidade mais antiga do Complexo do Caju, na zona portuária, foi uma das primeiras comunidades de baixa renda da cidade do Rio de Janeiro e do Brasil favorecidas pelo programa e aquela cujo processo de regularização se encontra mais avançado.

São poucos no mundo os trabalhos de avaliação das políticas públicas de regularização fundiária, assim como do impacto dos direitos de propriedade mais bem definidos e garantidos. Além disso, a literatura empírica sobre titularização de propriedades refere-se principalmente às áreas rurais. Existem alguns trabalhos, entretanto, que analisam programas de regularização de ocupações em comunidades de baixa renda em centros urbanos, como, por exemplo, Lanjouw e Levy (2002), no Equador; Field (2003a, 2003b, 2005) e Field e Torero (2006), no Peru; e Galiani e Schargrodsky (2004, 2007), na Argentina.

Os programas de regularização fundiária têm-se mostrado, internacionalmente, importantes ferramentas de combate à pobreza e à desigualdade. Dada a relevância que a academia e os formuladores de políticas conferem à regularização das ocupações irregulares para a melhoria das condições de vida da população de baixa renda, este trabalho pretende contribuir para o debate com uma avaliação do impacto do Programa de Regularização Fundiária da Prefeitura Municipal do Rio de Janeiro sobre o rendimento domiciliar *per capita* da Quinta do Caju, a partir de um painel com os mesmos domicílios da Quinta do Caju em 2002 (antes do programa) e 2008 (depois do programa).

Verificamos que o impacto do Programa de Regularização no rendimento domiciliar *per capita* dos moradores da Quinta do Caju foi positivo. Em especial, constata-se um aumento em torno de 20% a 32% na renda domiciliar *per capita* dos domicílios próprios tratados, em comparação com os não tratados. Esses resultados são robustos em relação a uma possível presença de viés de seleção pelas variáveis não observáveis, considerando a análise de sensibilidade de Rosenbaum.

O artigo está estruturado em cinco seções, além desta introdução. Na segunda seção é feita uma resenha da literatura empírica sobre a relação entre direitos de propriedade e bem-estar. Em seguida, são apresentados o Programa de Regularização Fundiária e a base de dados em painel utilizada no trabalho. Na quarta seção, abordamos a metodologia aplicada para avaliação do impacto do programa sobre o bem-estar. A quinta seção apresenta os resultados da avaliação de impacto. Por

1. O Favela-Bairro atua, desde 1995, em favelas de 500 a 2.500 domicílios da cidade do Rio de Janeiro, executando obras completas de urbanização, instalando creches, centros de assistência social e áreas de esporte e lazer, bem como realizando programas sociais e aprovando normas de controle urbanístico das áreas. Em especial, o programa teve cofinanciamento do Banco Interamericano de Desenvolvimento (BID) (VIAL; CAVALLIERI, 2009, p. 1).

fim, a conclusão sintetiza os principais pontos levantados ao longo do trabalho, com ênfase nos resultados obtidos e nas suas implicações para efeito de políticas públicas.

2 DIREITOS DE PROPRIEDADE E BEM-ESTAR: REVISÃO DA LITERATURA EMPÍRICA

Nos países e regiões onde os direitos de propriedade são bem definidos, atribuídos e assegurados, o que acontece como consequência? Diversos estudos analisaram, nos últimos anos, os efeitos desta situação.² Na literatura microeconômica, porém, o vínculo entre os direitos de propriedade e o aumento do bem-estar econômico tem, em geral, se limitado aos três canais estabelecidos por Besley (1995): *i*) aumento da segurança da ocupação e dos incentivos ao investimento; *ii*) custos de transação mais baixos e ganhos com o comércio; e *iii*) aumento do valor do ativo como garantia e diminuição da restrição ao crédito (FIELD, 2003a, p. 5).

Besley (1995) explorou os benefícios dos direitos de propriedade para os domicílios agrícolas em duas regiões rurais de Gana: Wassa e Anloga. Como Besley (1995), a maioria dos trabalhos empíricos do impacto dos direitos de propriedade formais refere-se à área rural, provavelmente devido ao interesse no investimento na agricultura e nas políticas de reforma agrária. Existem alguns trabalhos, entretanto, que analisam programas de regularização de ocupações em comunidades de baixa renda em centros urbanos como, por exemplo, Lanjouw e Levy (2002), Field (2003a, 2003b, 2005), Field e Torero (2006), Galiani e Schargrodsky (2004, 2007) e Payne, Durand-Lasserve e Rakodi (2007).³

Esta seção apresentará os efeitos econômicos, destacados na revisão da literatura, dos programas de titularização em áreas urbanas: *i*) investimento no domicílio/propriedade; *ii*) valor da propriedade; *iii*) acesso ao crédito; e *iv*) renda domiciliar e emprego.⁴

2.1 Investimentos na propriedade

Quanto ao primeiro canal estabelecido por Besley (1995), observa-se que as famílias, quando não possuem os direitos de propriedade de seus imóveis bem definidos e assegurados, tendem a investir menos na melhoria de seus domicílios com medo de perderem os investimentos. Dessa forma, um dos efeitos dos direitos de propriedade

2. Jimenez (1984) para Filipinas; Besley (1995) para Gana; Alston, Libecap e Schneider (1996) para o Norte do Brasil; Carter e Olinto (2000) para o Paraguai; Miceli, Sirmans e Kieyah (2001) para o Quênia; Lanjouw e Levy (2002) para o Equador; Do e Iyer (2002) para o Vietnã; Field (2003a, 2003b, 2005) e Field e Torero (2006) para o Peru; Galiani e Schargrodsky (2004, 2007) para a Argentina; Galeana (2004) para o México, entre outros.

3. Payne, Durand-Lasserve e Rakodi (2007) fizeram extensiva revisão da literatura sobre os impactos econômicos e sociais de 17 programas ou projetos que foram ou estão sendo implementados em áreas urbanas da África (Angola, Malawi, Senegal, África do Sul); Ásia (Afeganistão, Camboja, Índia, Indonésia, Laos, Filipinas) e América Latina (Argentina, Brasil, Bolívia, Colômbia, Equador, México e Peru).

4. Este artigo não focará nos efeitos sociais revisados por Payne, Durand-Lasserve e Rakodi (2007), como, por exemplo, *i*) segurança da ocupação; *ii*) status social, integração e inclusão social; *iii*) gênero; *iv*) saúde, educação e fertilidade; entre outros.

mais bem definidos e assegurados é o aumento da segurança na ocupação e, por conseguinte, dos investimentos duráveis na agricultura.

Assim como Besley (1995), Feder *et al.* (1988) na Tailândia, Alston, Libecap e Schneider (1996) no Norte do Brasil, Banerjee, Gertler e Ghatak (2002) em Bangladesh analisam o impacto dos títulos de propriedade no nível de investimento da agricultura. No entanto, Migot-Adholla *et al.* (1991), Place e Otsuka (2002) e Pinckney e Kimuyu (1994) observam pequeno impacto do título de propriedade no investimento.⁵

De acordo com Payne, Durand-Lasserve e Rakodi (2007), o fato de a obtenção do título de propriedade estimular o investimento no domicílio e a melhoria da propriedade tem sido o ponto-chave para justificar a implementação de programas de titularização. A hipótese básica é a de que os moradores só investiriam na melhoria do seu domicílio se a proteção da propriedade dos seus ativos e de seus investimentos estiver garantida por lei.

No caso de áreas urbanas, na ausência de direitos de propriedade bem definidos e assegurados, as famílias também tendem a gastar tempo extra e recursos para proteger seus imóveis de novas ocupações, diminuindo o investimento em bens duráveis e na melhoria dos seus domicílios.

No Peru, segundo Cantuarias e Delgado (2004, p. 9): *i*) 75% dos donos de domicílios com título de propriedade investiram na melhoria de suas casas, enquanto apenas 39% dos sem título fizeram investimentos; *ii*) os domicílios com título de propriedade têm maior número de quartos; e *iii*) os domicílios com título de propriedade são de melhor qualidade (paredes de tijolo, telhado de concreto, e com água).

Ainda quanto ao Peru, Field (2005) compara o investimento no domicílio antes e depois do programa entre os domicílios participantes e não participantes, e constata que os direitos de propriedade mais bem definidos e assegurados nas favelas urbanas têm um efeito significativo no investimento residencial. Field (2005) argumenta que, da mesma forma que as ocupações rurais, a insegurança da ocupação diminui os investimentos nos imóveis em regiões urbanas.

Por outro lado, Field (2005) aponta a possibilidade de um efeito negativo do título de propriedade no investimento no domicílio em áreas urbanas,⁶ o que a motivou a testar a relação e observar o efeito líquido do impacto do título no investimento no domicílio.

5. Esses resultados contraditórios são, em geral, atribuídos à dificuldade de resolver o problema de endogeneidade do *status* da ocupação (FIELD, 2005, p. 2).

6. Field (2005, p. 2-3) comenta, por exemplo, que os governos podem ser menos propensos a desapropriar comunidades cujos domicílios tenham suficientes investimentos.

Os dados de Field (2005) mostram que a taxa de melhoria do domicílio (*housing renovation*)⁷ cresce mais de dois terços (68%) do nível inicial apenas quatro anos após a obtenção do título de propriedade.⁸ Além disso, o estudo de Field (2005) sugere que o aumento do investimento está mais associado à diminuição da insegurança da ocupação do que com o aumento do acesso a crédito.

Similarmente, Galiani e Schargrodsky (2004, 2007) analisam o impacto dos títulos de propriedade em comunidades de baixa renda dos subúrbios de Buenos Aires, na Argentina. Os resultados de Galiani e Schargrodsky (2007) apresentam efeitos significativos em investimentos no domicílio, assim como no tamanho das famílias e na frequência do ensino secundário.

Quanto aos efeitos do título de propriedade nos investimentos no domicílio, Galiani e Schargrodsky (2007, p. 16-17) apontam quatro canais dessa relação: *i*) aumento da segurança da ocupação; *ii*) aumento da liquidez dos investimentos; *iii*) acesso ao mercado de crédito; e *iv*) instrumento de poupança.⁹

Galiani e Schargrodsky (2007) estabeleceram um índice para a qualidade total do imóvel,¹⁰ que varia de 0 a 100 e sintetiza todos os aspectos do imóvel. Em particular, em relação aos valores da linha de base, o efeito estimado do título na melhoria do imóvel é de 37%.

Por outro lado, Payne, Durand-Lasserve e Rakodi (2007, p. 14-15) ressaltam que a má definição e a falta de garantia dos direitos de propriedade não impedem os investimentos em infraestrutura. Em países onde a ameaça ou a percepção de remoção é mínima ou inexistente, como Turquia, Trinidad, Egito, Marrocos e cidades da África Ocidental, proprietários estão dispostos a investir com ou sem títulos de propriedade. Dessa forma, Payne, Durand-Lasserve e Rakodi (2007, p. 14) apontam que a percepção da segurança e o alcance de uma mínima cesta de direitos de propriedade exercem mais influência sobre o nível de investimento e outros benefícios do que o título de propriedade em si.

Além disso, outro ponto levantado pelos autores é que o nível baixo de investimento no domicílio pode estar mais associado à falta de renda do que à má definição e falta de garantia dos direitos de propriedade.

7. Field (2005) considerou as seguintes variáveis para melhoria do domicílio: *i*) paredes novas; *ii*) novo telhado; *iii*) paredes pintadas; *iv*) eletricidade; *v*) acesso a água; *vi*) banheiro adicional; *vii*) quarto adicional; entre outras.

8. Em particular, a natureza dos investimentos está mais relacionada a pequenas melhorias na residência do que com expansão do imóvel.

9. O título de propriedade permite ao proprietário substituir o consumo e o lazer presentes por poupança de longo prazo em bens imóveis.

10. Este índice foi construído por uma equipe de arquitetos que foi até o local para avaliar as condições de cada domicílio da amostra (GALIANI; SCHARGRODSKY, 2007).

Por exemplo, no México, Angel *et al.* (2006, p. 14) constataram que, devido ao fato de a propriedade ser relativamente segura mesmo sem título, os proprietários não esperam ter a escritura de seus imóveis para fazer melhorias. Em especial, a baixa taxa de investimento no imóvel está mais relacionada com a renda limitada dos proprietários do que com a falta de título de propriedade.

Portanto, podemos concluir, a partir dos trabalhos empíricos discutidos acima, que o efeito do título no investimento e na melhoria do domicílio é positivo, devido ao aumento de segurança proporcionado pelo título. No entanto, a magnitude desse aumento depende do grau de segurança da ocupação antes do programa de titularização. No caso das favelas cariocas, por exemplo, podemos esperar que, como estas já estão bem consolidadas e as políticas públicas têm sido nos últimos anos mais focadas na melhoria das condições habitacionais dessas comunidades, o efeito do título no investimento do imóvel deva ser relativamente pequeno.

2.2 Valor da propriedade

O segundo custo decorrente dos direitos de propriedade mal definidos e assegurados, como apresentado por Besley (1995), é a limitação dos ganhos advindos do comércio. Com os direitos de propriedade mais bem definidos e assegurados, há aumento das possibilidades de ganho com a troca e, por conseguinte, do valor de mercado do ativo em questão. Os direitos de propriedade formalmente reconhecidos permitem o comércio com estranhos e não somente com um círculo restrito de conhecidos. Alston, Libecap e Schneider (1996, p. 29) constataram, no Norte do Brasil, que o título de propriedade é percebido como uma vantagem pelos ocupantes, pois amplia o leque de possíveis compradores (ver, também, JIMENEZ, 1984; FRIEDMAN; JIMENEZ; MAYO, 1988; ALSTON; LIBECAP; SCHNEIDER, 1996; LANJOUW; LEVY, 2002).

Lanjouw e Levy (2002) mostram que o preço de mercado de uma propriedade reflete tanto a utilidade do proprietário do ativo quanto a confiança do comprador de que a compra seja honrada. Quando o comprador não está certo de que a família irá honrar a venda da sua propriedade ou quando o proprietário não está confiante de que o inquilino cumprirá o acordo, o conjunto de pessoas com quem as famílias podem transacionar é mais limitado, talvez se restrinja somente aos amigos ou familiares. Em geral, essa limitação impede a alocação da propriedade para famílias que a valorizariam mais (LANJOUW; LEVY, 2002, p. 1.011). Portanto, quando os direitos de propriedade são formalizados, devem-se interpretar as mudanças no preço como mudanças na utilidade do proprietário assim como no grau de incerteza da transação (LANJOUW; LEVY, 2002, p. 987).

Segundo os autores, o título de propriedade proporciona a segurança da ocupação e, por conseguinte, permite também às famílias captarem os investimentos feitos em ativos duráveis no domicílio. Além disso, os títulos de propriedade,

tornando os direitos mais claros e transferíveis, reduzem os custos de transação no mercado. Mesmo para as famílias que não têm intenção de vender seus imóveis, a possibilidade de alienar a propriedade é válida, pois permite o uso da propriedade como garantia para empréstimos formais (LANJOUW; LEVY, 2002, p. 988).

No entanto, Lanjow e Levy (2002, p. 987) apontam que a possibilidade de transferência do imóvel também é importante quando se analisam os impactos dos programas de titularização.¹¹ Com direitos mais bem definidos, contudo sem a possibilidade de transferência da propriedade, a transação do imóvel pode ser mais difícil e incerta (LANJOUW; LEVY, 2002, p. 987).

Uma das principais contribuições do trabalho de Lanjouw e Levy (2002) foi a relação feita entre os direitos de propriedade formais e informais. Os autores observam que o efeito dos direitos de propriedade formais no comportamento e bem-estar econômico depende da presença de fontes informais de direitos de propriedade, e que estas últimas são também importantes na estimativa do valor das propriedades com direitos formais (LANJOUW; LEVY, 2002, p. 986). Em especial, o título de propriedade tem efeito maior sobre o bem-estar quando as relações informais estabelecidas na comunidade são fracas.

Dessa forma, no estudo empírico de uma comunidade de baixa renda em área urbana (cidade de Guayaquil) no Equador, os autores mostram que o impacto dos direitos de propriedade formais sobre o preço dos imóveis é maior em recentes ocupações sem organização, onde os direitos de propriedade informais são fracos. O preço de mercado esperado para as propriedades com título, em comunidades com somente um ano de existência e sem organização, aumenta em 51,6% em relação às propriedades sem os direitos de propriedade formais, mas com as mesmas características (LANJOUW; LEVY, 2002, p. 989). Enquanto, em geral, ocorre um aumento de 23,5% no valor das propriedades quando existem direitos de propriedade informais (LANJOUW; LEVY, 2002, p. 989).

No Peru, Cantuarias e Delgado (2004, p. 9) constatam, a partir dos dados coletados pela Comisión de Formalización de la Propiedad Informal (Cofopri), que o impacto do título de propriedade no valor do imóvel é de aproximadamente 25%. Esta estimativa é confirmada no estudo de Angel *et al.* (2006, p. 11). Adicionalmente, de acordo com os dados da linha de base da Cofopri, 64% das famílias que receberam o título de propriedade no Peru acreditam que o título aumenta o valor dos seus imóveis (CANTUARIAS; DELGADO, 2004, p. 9).

11. Vale notar que alguns programas de titularização concedem títulos de propriedade para os ocupantes, mas fazem ressalvas quanto à venda ou transferência desses imóveis. Ver, por exemplo, o caso do programa de titularização em Buenos Aires, analisado por Galiani e Scharrodsky (2007), que concedeu os títulos mas estabeleceu que os ocupantes não poderiam transferir a propriedade nos dez anos seguintes ao do recebimento do título.

Portanto, a literatura internacional constata que um dos efeitos da titularização é o aumento do valor dos imóveis. No entanto, é importante salientar que esse efeito é maior quando as comunidades beneficiadas pelo programa são recentes e sem organização, ou seja, quando as instituições informais são fracas.

2.3 Acesso ao crédito

O terceiro custo decorrente dos direitos de propriedade mal definidos e assegurados, como apresentado por Besley (1995), refere-se ao argumento principal de De Soto (2000): restrição ao crédito das comunidades pobres. Os imóveis sem direitos de propriedade bem definidos também não servem como garantia para empréstimos, impedindo o acesso dos pobres ao mercado de crédito. Vários trabalhos analisam o efeito do título de propriedade sobre o mercado de crédito (FEDER *et al.*, 1988; LOPEZ; ROMANO, 1997; MIGOT-ADHOLLA *et al.*, 1991; CARTER; WIEBE; BLAREL, 1994; ALSTON; LIBECAP; SCHNEIDER, 1996; CARTER; OLINTO, 2000; GALIANI; SCHARGRODSKY, 2007; FIELD; TORERO, 2006; GALEANA, 2004), porém os resultados não são os mesmos. A maioria deles confirma, em parte, a hipótese de De Soto (2000) de que a formalização dos direitos de propriedade diminui a restrição ao crédito ao proporcionar garantia para empréstimos.

Os estudos empíricos de Feder *et al.* (1988) na Tailândia, e de Lopez e Romano (1997) em Honduras indicaram um efeito positivo da regularização fundiária de áreas rurais sobre o acesso ao crédito. Por outro lado, Migot-Adholla *et al.* (1991) em Gana, Quênia e Ruanda; Carter, Wiebe e Blarel (1994) no Quênia; Carter e Olinto (2000) no Paraguai; Galiani e Schargrotsky (2007) na Argentina; Field e Torero (2006) no Peru; e Galeana (2004) no México não observam um efeito muito significativo no acesso ao crédito dos proprietários com título. Migot-Adholla *et al.* (1991) argumentam que o resultado negativo é devido à falta de um efetivo sistema de crédito nesses países. Já Carter e Olinto (2000) mostram que a oferta de crédito só aumentou para os médios e grandes fazendeiros, não para os pequenos fazendeiros, depois do recebimento do título da terra no Paraguai.

Em especial, os artigos de Galiani e Schargrotsky (2007) para uma área do subúrbio de Buenos Aires, e de Field e Torero (2006), para áreas urbanas do Peru são alguns dos poucos trabalhos que analisam a relação entre direitos de propriedade e acesso ao mercado de crédito em regiões urbanas.

Os resultados de Galiani e Schargrotsky (2007) constataram um impacto muito pouco significativo do título de propriedade no acesso ao mercado de crédito formal. Ao comparar as consequências do título de propriedade em duas regiões idênticas, sendo uma com título e outra sem título, os autores não constataram diferença alguma entre os dois grupos no acesso a cartão de crédito, conta-corrente em bancos, ou empréstimos de bancos, governo ou cooperativas. No entanto, essas

famílias têm baixo acesso ao crédito formal. O acesso ao crédito, como constatam os autores, é maior para o crédito informal, concedido por parentes, colegas de trabalho, vizinhos, amigos ou estabelecimentos em que compram com frequência. Contudo, os autores não observam o efeito do título no acesso a essas fontes informais de crédito. Em relação ao mercado de crédito formal, Galiani e Scharfrodsky (2007, p. 27) observam que os domicílios sem título não receberam financiamento, enquanto apenas 4% dos domicílios com título receberam empréstimos formais.

Os resultados de Field e Torero (2006) indicam a permanência de racionamento no mercado de crédito formal no Peru, apesar da distribuição de 1,2 milhão de títulos de propriedade. Depois dessa distribuição, 34% dos proprietários que receberam títulos de propriedade mantiveram-se restritos ao crédito formal. Entretanto, a maior parte da redução no racionamento do crédito, associada aos títulos de propriedade urbana, pode ser atribuída, segundo os autores, à instituição pública de empréstimo para a compra de material de construção. Em particular, houve um aumento de 10% na taxa de aprovação de crédito formal obtido em bancos públicos e uma redução de 9 pontos na taxa de juros para crédito em bancos privados (FIELD; TORERO, 2006).

Em suma, na literatura internacional observa-se que o efeito do título de propriedade sobre o crédito ainda é pouco significativo, e muitas vezes os resultados não são conclusivos. O efeito causal positivo dos direitos de propriedade no crédito, como defende De Soto (2000), depende também do estágio de desenvolvimento do setor financeiro e bancário, assim como do sistema jurídico. Por exemplo, no Brasil, é importante destacar que o imóvel utilizado como residência da família não pode ser utilizado como instrumento de garantia de crédito.

2.4 Rendimento domiciliar e emprego

Outro efeito econômico dos direitos de propriedade mais bem definidos e assegurados, apontado na literatura, seria o aumento da renda domiciliar e do emprego.

A necessidade permanente de ter um adulto em casa para proteger o domicílio contra nova ocupação reduz a oferta de trabalho dos adultos e aumenta a oferta de trabalho infantil. Segundo Field (2003a), em seu estudo de uma comunidade de baixa renda urbana do Peru, a principal consequência da formalização dos direitos de propriedade é a redução da probabilidade de as famílias serem expulsas de seus domicílios.

Dessa forma, a autora constata que, depois do programa de titularização, há o aumento do número total de horas de trabalho e a realocação das horas de trabalho de dentro para fora do domicílio. Em particular, Field (2003a) observa que as famílias com título de propriedade de seus imóveis trabalham em média 17% de horas a mais do que as famílias que aguardam receber o título de propriedade,

e 38% das famílias com título têm uma probabilidade maior de desenvolverem atividades fora de casa. Além disso, observa também uma redução em 47% da probabilidade do trabalho dentro de casa e, nas famílias com criança, uma diminuição de 28% da probabilidade de participação infantil no trabalho. Além disso, Field (2003a) constata que direitos de propriedade informal e tamanho da família influenciam no grau de segurança da ocupação.

No entanto, como apontaram Payne, Durand-Lasserve e Rakodi (2007), o trabalho de Field (2003a) não esclarece se o aumento de horas de trabalho é seguido por aumento de salários.¹²

Em relação ao rendimento domiciliar, Andrade (2004) verificou econometricamente que a formalização dos direitos de propriedade apresenta impacto positivo e significativo no rendimento domiciliar *per capita* esperado nas comunidades do Caju, na cidade do Rio de Janeiro, a partir da Pesquisa Socioeconômica das Comunidades de Baixa Renda do Caju de 2002.¹³ Quando controlado pelas variáveis de riqueza do domicílio, região do complexo, características pessoais, de escolaridade e de ocupação do chefe de domicílio, de escolaridade e trabalho do cônjuge e dos outros moradores do domicílio, o rendimento domiciliar *per capita* esperado dos domicílios com escritura definitiva dessas comunidades é 15% superior em relação aos domicílios sem nenhuma definição e garantia de seus direitos de propriedade.

No entanto, Galiani e Schargrotsky (2007), no seu estudo para um bairro de Buenos Aires, não constataram impacto significativo do título de propriedade na renda do chefe do domicílio, renda domiciliar total, renda domiciliar *per capita*, renda domiciliar total dos adultos, e *status* ocupacional do chefe do domicílio. Além disso, também não observaram nenhuma diferença significativa no emprego feminino ou no trabalho infantil. Apesar do título de propriedade, os autores observam que as famílias continuam sendo muito pobres após 20 anos do recebimento do título: a renda familiar média equivale a apenas 38% da linha oficial de pobreza, e 94% dos domicílios estão abaixo dessa linha (GALIANI; SCHARGRODSKY, 2007, p. 29).

Um ponto levantado por Fernandes (2006, *apud* PAYNE; DURAND-LASSERVE; RAKODI, 2007) é que os moradores de favelas, mesmo depois de receberem seus títulos de propriedade, ainda são vistos – e eles próprios se

12. De acordo com Payne, Durand-Lasserve e Rakodi (2007, p. 21), muitas das áreas beneficiadas por títulos de propriedade nos arredores de Lima, e outras cidades grandes do Peru, situam-se longe das áreas de maior emprego e, portanto, o tempo e o custo adicional para se locomover até esses locais podem não representar um benefício significativo para essas famílias.

13. Vale notar, no entanto, que a análise de Andrade (2004) foi feita para todas as comunidades de baixa renda do Caju antes do Programa de Regularização Fundiária na Quinta do Caju. Dessa forma, a autora comparou os domicílios que declararam ter escritura de seus imóveis com os domicílios sem os direitos de propriedade bem definidos e assegurados (com apenas documentação da associação de moradores ou nenhum tipo de documentação).

sentem – como favelados e, portanto, discriminados no mercado de trabalho.¹⁴ Dessa forma, o título de propriedade pode não implicar um aumento da renda do trabalho desses moradores.

Portanto, observa-se que, em geral, o vínculo entre direitos de propriedade e renda foi testado empiricamente por meio do canal do mercado de trabalho, ou seja, a hipótese testada foi sempre a de que o título aumenta a segurança da ocupação e, por conseguinte, possibilita um aumento da oferta de trabalho. Contudo, vale ressaltar que os direitos de propriedade podem influenciar a renda domiciliar tanto pelo canal do crédito quanto da riqueza. O aumento do valor da propriedade, seja pela apropriação dos investimentos no domicílio ou pela diminuição dos custos de transação, possibilita um acesso maior ao crédito, ao utilizar a propriedade como garantia, e aumento das possibilidades de transação com o imóvel (venda ou aluguel).

3 O PROGRAMA E A BASE DE DADOS

O Caju, localizado na zona portuária do centro do Rio de Janeiro, é um dos bairros mais antigos da cidade e, além disso, apresenta muitas especificidades devido ao seu processo histórico de ocupação. Há nove comunidades de baixa renda espalhadas no bairro do Caju,¹⁵ entre fábricas, áreas militares, estaleiros, depósito da Companhia Municipal de Limpeza Urbana (COMLURB) etc.

A Quinta do Caju é a comunidade mais antiga. Sua ocupação por pescadores, na maioria imigrantes portugueses, começou provavelmente na segunda metade do século XIX. A comunidade é, por esse motivo, bastante homogênea (formada por idosos) e com pouca mobilidade (os moradores moram em média há 40 anos no mesmo domicílio). Além disso, a comunidade é bem organizada, com uma associação de moradores muito ativa e politicamente participativa, que intermedeia internamente suas relações comerciais. As relações pessoais e as instituições informais são fortes na comunidade. Além disso, sua ocupação é segura e consolidada na cidade do Rio de Janeiro.

Nesta seção, serão apresentados o Programa de Regularização Fundiária e a base de dados utilizada neste trabalho.

3.1 O Programa de Regularização Fundiária na Quinta do Caju

A região hoje conhecida como Quinta do Caju foi doada a D. João VI, rei de Portugal, em 1815, pelo proprietário Luís José Gouveia e Freire. Com a Proclamação da República, essas terras passaram a integrar o patrimônio da União. No final do

14. Para uma análise sobre discriminação salarial contra moradores em favelas, ver Cardoso, Elias e Pero (2003).

15. As nove comunidades que formam o Complexo do Caju são: Ladeira dos Funcionários, Parque São Sebastião, Clemente Ferreira, Quinta do Caju, Parque Alegria, Parque Vitória, Parque Conquista, Nossa Senhora da Penha e Parque Boa Esperança.

século XIX, surgiu a primeira colônia do bairro ocupada por pescadores, em sua maioria, descendentes de portugueses.

Devido à precariedade das condições da Quinta do Caju, a prefeitura implementou alguns programas habitacionais na comunidade, a começar pelo Programa Favela-Bairro e, posteriormente, com a regularização fundiária.

A Prefeitura do Rio de Janeiro, por intermédio da Secretaria Municipal de Habitação, iniciou o processo de regularização das moradias no Complexo do Caju, pela Quinta do Caju, em 1998. Trata-se de um programa pioneiro na medida em que o terreno da Quinta do Caju pertencia à União. Dessa forma, o processo de regularização fundiária na comunidade deu-se por um contrato em que a União cedia a área ao município sob regime de aforamento,¹⁶ para que o município pudesse vender, em condições facilitadas, os terrenos a seus ocupantes, e fazer investimentos de infraestrutura no local.¹⁷

Vale notar que a Quinta do Caju foi beneficiada pelo Programa de Regularização Fundiária devido também à pressão feita pela associação de moradores junto às autoridades locais e federais.

Contudo, apesar de o processo de regularização fundiária na Quinta do Caju ter-se iniciado em 1998, ele foi concluído junto à União apenas em dezembro de 2004. Dessa forma, os primeiros domicílios receberam o termo de compra e venda, registrado em Cartório de Registro de Imóveis, somente em junho de 2004.

Segundo dados da prefeitura, no final do primeiro semestre de 2008, dos 843 domicílios da Quinta do Caju, 22,3% deles receberam a escritura definitiva dos seus imóveis e 10,3% receberam a declaração de Concessão de Direito Real de Uso (CDRU), que é uma permissão para registrar as propriedades em Cartório de Registro de Imóveis. Os demais domicílios ainda estão em processo de regularização junto à prefeitura, como mostra a tabela 1.

A Coordenação de Regularização Fundiária da Secretaria de Habitação da Prefeitura do Rio de Janeiro é responsável pelo Programa de Regularização Fundiária nas favelas cariocas.¹⁸ Segundo essa Coordenação, no início do Programa

16. Aforamento é o contrato pelo qual o proprietário atribui a alguém o domínio útil de um imóvel, mediante pagamento anual de uma quantia, denominada foro.

17. No caso da Quinta do Caju, a União, que era proprietária do terreno, passou o domínio útil para o município por meio de um contrato de cessão e este último o venderá para os moradores, que registrarão o termo de compra e venda do imóvel na Gerência Regional do Patrimônio da União e no Registro Geral de Imóveis. Dessa forma, os moradores passarão a ser coproprietários de seus lotes, com a União Federal, que continuará a possuir o domínio direto. No caso dos moradores da Quinta do Caju, eles estão isentos do pagamento de qualquer taxa à União, no caso, do foro (anual), já que se trata de comunidades de baixa renda que estão isentas desse pagamento. No entanto, eles podem usufruir o imóvel de forma mais completa possível, inclusive transferindo a seus herdeiros.

18. Em 11 de julho de 2008 foi feita uma entrevista com a coordenadora responsável pelo Programa de Regularização Fundiária, Angélica Pullig, que explicou o processo de regularização fundiária na Quinta do Caju e forneceu informações sobre o programa.

de Regularização Fundiária, a prefeitura cadastrou todos os domicílios e fez as medições de cada imóvel (lote) da Quinta do Caju. Dessa forma, os moradores ficaram cientes do programa e, a partir de 2004, houve duas entregas conjuntas dos termos de compra e venda na comunidade.

TABELA 1

Programa de Regularização Fundiária na Quinta do Caju

Ano	CRDU	Termo de compra e venda
2004	44	42
2005	35	100
2006	4	35
2007	3	14
1º semestre de 2008	1	2
Total	87	193

Fonte: Coordenação de Regularização Fundiária, Secretaria de Habitação, Prefeitura do Rio de Janeiro.

Para dar entrada ao pedido de regularização do imóvel junto à prefeitura são necessários os seguintes documentos: identidade, Cadastro de Pessoas Físicas (CPF), comprovante de residência, certidão de casamento e documentos do cônjuge. No caso de viúvos, certidão de óbito do cônjuge. Não é obrigatório apresentar comprovante de renda.

Os moradores devem entregar uma cópia destes documentos na prefeitura. Vale apontar que a associação de moradores é atuante na comunidade ao prestar informações sobre o programa e levar a documentação à prefeitura.

Quanto ao custo da regularização, para emissão do termo de compra e venda, é necessário o pagamento de uma taxa no valor de R\$ 2,00/m² para um único imóvel do titular, R\$ 13,46/m² (gleba “A”) e R\$ 19,23/m² (gleba “B”) para os demais imóveis do titular e para as unidades comerciais.¹⁹

O valor da taxa para obtenção do termo de compra e venda pode ser dividido em até 60 parcelas. Caso o proprietário opte pelo pagamento parcelado da taxa, é assinado um contrato de CDRU com o número e o valor das parcelas. Dessa forma, é enviado um carnê para a residência do proprietário com os boletos de pagamento.

Caso o proprietário opte pelo pagamento à vista, é feita uma ficha de depósito para o pagamento total da taxa. E, em seguida, é encaminhado o processo para o Cartório de Registro de Imóveis, a fim de ser registrado o termo de compra e venda. A própria prefeitura se encarrega deste procedimento, e não há custo para o proprietário.

19. Nesta primeira transferência da propriedade aos moradores não é cobrado o laudêmio que recairia sobre a transação, por se tratar de terreno da Marinha. Também não é exigido o Imposto sobre a Transmissão de Bens Imóveis (ITBI), que seria pago pelo comprador ao município.

Entretanto, antes de encaminhar para o Cartório de Registro de Imóvel, a prefeitura precisa obter uma certidão autorizativa de transferência junto à Gerência Regional do Patrimônio da União no Rio de Janeiro.

Portanto, de posse de toda a documentação e da certidão autorizativa de transferência, a prefeitura encaminha o processo para o Cartório de Registro de Imóveis. O processo junto ao cartório leva mais ou menos 30 dias, se não for feita nenhuma exigência quanto à documentação. No entanto, a Coordenação de Regularização Fundiária alerta que pode surgir problema quanto à planta do lote, quando se trata de lotes desmembrados, ou seja, quando um imóvel se desmembra em mais de um domicílio.

Além disso, a coordenadora do programa diz ser comum os moradores procurarem a prefeitura pedindo informações sobre a regularização fundiária, mas apresentarem problemas de documentação.²⁰ Vale destacar que o morador não pode ser beneficiado pelo programa antes de apresentar toda a documentação necessária.

A Coordenadoria de Regularização Urbanística é responsável pela emissão da Certidão de Habite-se. Esta certidão oficializa o número do imóvel, que será reconhecido pelos demais órgãos do município e perante o Cartório de Registro de Imóveis. Na certidão, constam o endereço oficial, o número de pavimentos e a metragem quadrada construída.

Após o recebimento do título, a Coordenadoria de Imposto Predial e Territorial Urbano (IPTU) da prefeitura deve incluir o imóvel no cadastro fiscal. As informações são geradas com base no formulário de Habite-se encaminhado pela Secretaria de Habitação e Urbanismo. No entanto, a inclusão do imóvel no cadastro do IPTU não acarreta obrigatoriamente a cobrança do imposto.

3.2 Base de dados

A base de dados utilizada neste trabalho é um painel construído a partir da Pesquisa Socioeconômica das Comunidades de Baixa Renda do Caju (PSCBRC), elaborada em setembro de 2002 pelo Instituto de Estudos do Trabalho e Sociedade (IETS) em parceria com o sistema Federação das Indústrias do Estado do Rio de Janeiro (Firjan) e a Sociedade Científica da Escola Nacional de Ciências Estatísticas (Science), e a Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em agosto de 2008, com financiamento da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ). O painel utilizado refere-se aos mesmos domicílios²¹ da Quinta do Caju em 2002 (antes do programa) e 2008 (depois do programa).

20. Como, por exemplo: *i*) idosos sem documentação e com dificuldades de providenciar a documentação requerida; *ii*) mulher, cujo marido desapareceu, não tem mais a documentação do cônjuge ou não quer que ele tenha direito ao imóvel; e *iii*) problema de inventário e briga entre os filhos.

21. Trata-se dos mesmos domicílios, pois a pesquisa de campo de 2008 levou em conta a listagem amostral de 2002. Dessa forma, foram selecionados os mesmos domicílios da Quinta do Caju da pesquisa PSCBRC de 2002 para reentrevistar os moradores.

Vale ressaltar que, para ambas as pesquisas, foi aplicado praticamente o mesmo questionário. Portanto, para a construção do painel, as variáveis referentes às características dos domicílios e seus moradores são as mesmas para os dois anos.

Apesar de ter sido tomada por base uma amostra de 261 domicílios particulares permanentes da Quinta do Caju que fizeram parte da pesquisa PSCBRC de 2002, foram reentrevistados 231 domicílios em 2008. Houve um atrito de 30 domicílios, ou seja, uma perda de aproximadamente 13% dos domicílios da amostra inicial de 2002. Dos 30 domicílios não entrevistados, 17 eram domicílios vagos, 3 domicílios de uso ocasional, 7 domicílios fechados e 3 em que houve recusa. Além disso, dos 231 domicílios entrevistados em 2008, não encontramos informações referentes ao ano de 2002 para dois domicílios.²² Dessa forma, o painel só pôde ser construído com uma amostra de 229 domicílios particulares permanentes da Quinta do Caju.

A Pesquisa Socioeconômica da Comunidade da Quinta do Caju de 2008 foi aplicada nos domicílios particulares permanentes da amostra da Quinta do Caju da PSCBRC de 2002. Mas, no universo dos domicílios particulares permanentes, nos restringimos apenas aos domicílios que eram próprios²³ tanto em 2002 quanto em 2008, já que os direitos de propriedade são o objeto deste trabalho. Dessa forma, foram excluídos da análise os domicílios alugados, e os que foram cedidos por empregador ou de outra forma. Com essa seleção, a amostra apresenta 185 domicílios particulares permanentes próprios.

Os domicílios próprios diferem quanto à condição de ocupação, ou seja, em relação à definição de seus direitos de propriedade. Os domicílios próprios podem apresentar o termo de compra e venda, registrado em Cartório de Registro de Imóveis, declaração de CDRU emitida pela prefeitura, documentação apenas da associação dos moradores, nenhuma das documentações anteriores ou que ainda não acabou de pagar.

Além disso, dos domicílios da amostra, 89,7% referem-se à mesma família, ou seja, foram reentrevistadas, na maior parte da amostra, as mesmas famílias de 2002. No entanto, das mesmas famílias, 37,3% continuam com os mesmos moradores e 52,4% têm moradores a mais ou a menos na família. Adicionalmente, dos domicílios da amostra, 74% referem-se ao mesmo chefe do domicílio.

Dessa forma, foi possível construir três painéis diferentes: *i*) domicílios próprios – com 185 observações; *ii*) domicílios próprios e mesma família – com 166 observações; e *iii*) domicílios próprios e mesmo chefe – com 137 observações.

22. Provavelmente houve perda de informação quando a base de dados de 2002 foi digitada.

23. Os domicílios particulares permanentes são classificados como próprios quando o domicílio é de propriedade, total ou parcial, do morador, estando integralmente quitado ou não, e independentemente da condição de ocupação do terreno.

3.2.1 Grupos de tratamento e de comparação

Como nosso objetivo é avaliar o impacto do Programa de Regularização Fundiária na Quinta do Caju, precisamos construir um grupo que foi beneficiado pelo programa (grupo de tratamento) e outro grupo que não foi beneficiado (grupo de comparação). Além disso, como o Programa de Regularização Fundiária na Quinta do Caju começou a ser implementado em 2004, entre as duas pesquisas de campo, teremos dados desses dois grupos antes e após a implementação inicial do programa.

Definiremos, portanto, como grupo de tratamento os domicílios que foram beneficiados pelo Programa de Regularização Fundiária, ou seja, os domicílios que têm o termo de compra e venda (TCV), registrado em Cartório de Registro de Imóveis, ou a declaração de CDRU emitida pela prefeitura. Enquanto isso, definiremos como grupo de comparação os demais domicílios da Quinta do Caju, que ainda não foram beneficiados pelo programa (não apresentam TCV nem CDRU).

Dessa forma, da amostra de domicílios próprios, constata-se que 39,5% pertencem ao grupo de tratamento, e o restante (60,5%) ao grupo de comparação, como apresentado na tabela 2. Vale ainda notar que, dos domicílios do grupo de tratamento, 65,8% apresentam o TCV, e 34,2% possuem a declaração de CDRU.

TABELA 2

Quinta do Caju: grupos de tratamento e comparação

Grupos	Observações	%
Grupo de tratamento	73	39,5
Grupo de comparação	112	60,5
Total	185	100,0

Fontes: PSCBRC, elaborada em 2002 pelo Iets/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

Um ponto a ser questionado é se o grupo de comparação é um bom controle para o grupo de tratamento, uma vez que o definimos por exclusão, considerando os domicílios que não foram beneficiados pelo programa. Como o grupo de comparação não foi definido aleatoriamente, os dois grupos podem não ser parecidos tanto por características observáveis quanto não observáveis, o que pode causar um viés de autosseleção na avaliação do impacto do programa. Além disso, como já vimos, para participar do programa, os próprios moradores se apresentam à prefeitura para pagar pelo seu título de propriedade. Portanto, a seleção ao programa é feita pelos próprios moradores, não de forma aleatória ou experimental.

Para investigar se os domicílios do grupo de tratamento e comparação são parecidos, observaremos se existiam diferenças estatisticamente significativas quanto às características desses domicílios e de seus moradores em 2002, antes do Programa de Regularização Fundiária ter sido implementado pela prefeitura na Quinta do

Caju. Portanto, na próxima subseção, apresentaremos o teste das diferenças de média entre os grupos de tratamento e de comparação antes do programa.

3.2.2 Teste das diferenças de média entre os grupos antes do programa

A tabela 3 revela que as diferenças entre os grupos de tratamento e de comparação, antes da realização do programa, não são estatisticamente significativas para a maioria das características do domicílio, dos chefes e dos cônjuges. As diferenças estatisticamente significativas entre os dois grupos são quanto às variáveis que indicam a presença de crianças com 5 a 11 anos, o número de adultos no domicílio, o número de cômodos, o acesso à água e a idade do chefe e do cônjuge.

Primeiramente, constata-se que os domicílios do grupo de tratamento têm, em média, menos crianças e mais adultos. Em segundo lugar, os domicílios do grupo de tratamento têm, em média, um número maior de cômodos e menos acesso à água. Por fim, quanto às características dos moradores, os chefes e cônjuges dos domicílios do grupo de tratamento são, em média, mais velhos.

A diferença do rendimento domiciliar *per capita* entre os dois grupos antes da realização do programa não era estatisticamente significativa. Além disso, considerando as variáveis se o chefe/cônjuge tem identidade e CPF, que são os documentos necessários para a inscrição no programa, as diferenças não são estatisticamente significativas entre os dois grupos. Da mesma forma, a diferença de escolaridade e do nível de ocupação do chefe e do cônjuge entre os dois grupos não é estatisticamente significativa.

Conclui-se que os grupos de tratamento e de comparação, em relação às características observáveis dos domicílios e moradores, não apresentavam diferenças estatisticamente significativas muito importantes antes da realização do programa e, portanto, o grupo de comparação pode servir de controle para o tratamento quando for feita a avaliação de impacto do Programa de Regularização na Quinta do Caju.

TABELA 3

Teste das diferenças entre grupos de tratamento e de comparação – 2002

Pré-programa (2002)	Grupo de tratamento	Grupo de comparação	Diferença	Valor-p
Composição do domicílio				
Número de moradores	2,95	3,29	0,35	0,16
Número de moradores em idade ativa	2,74	2,98	0,24	0,23
Número de crianças com 0 a 5 anos de idade	0,11	0,17	0,06	0,43
Número de crianças com 5 a 11 anos de idade	0,07	0,24	0,17	0,01**
Número de crianças com 12 a 14 anos de idade	0,07	0,15	0,08	0,10

(continua)

(continuação)

Pré-programa (2002)	Grupo de tratamento	Grupo de comparação	Diferença	Valor-p
Número de adolescentes com 15 a 17 anos de idade	0,36	0,33	-0,03	0,82
Número de jovens (18 a 24 anos)	1,37	1,62	0,25	0,11
Número de adultos (24 a 60 anos)	0,86	0,57	-0,29	0,01**
Número de idosos (60 anos ou mais)	0,11	0,21	0,10	0,12
Características do domicílio				
Anos na comunidade	39,67	39,68	0,01	1,00
Paredes de alvenaria/tijolo com revestimento	94,5%	87,5%	-7,0%	0,12
Telhado (telha ou laje)	95,9%	94,6%	-1,2%	0,70
Piso (madeira, carpete, cerâmica, cimento)	97,3%	96,4%	-0,9%	0,75
Quantos cômodos no domicílio	5,49	5,05	-0,44	0,04**
Quantos cômodos servem de dormitórios	1,64	1,84	0,20	0,21
Acesso adequado à água	95,9%	100,0%	4,1%	0,03**
Esgotamento sanitário adequado	100,0%	100,0%	0,0%	
Lixo é coletado diretamente	100,0%	100,0%	0,0%	
Acesso à eletricidade formal	98,6%	99,1%	0,5%	0,76
Acesso a fogão	94,5%	96,4%	1,9%	0,54
Acesso à geladeira	100,0%	98,2%	-1,8%	0,25
Acesso à televisão	98,6%	99,1%	0,5%	0,76
Acesso a telefone (fixo ou celular)	90,4%	87,5%	-2,9%	0,54
Acesso a computador	13,7%	14,3%	0,6%	0,91
Características do chefe				
Chefe é mulher	28,8%	33,0	4,3	0,54
Idade do chefe	60,97	56,90	-4,07	0,03**
Chefe é alfabetizado	95,9%	90,2%	-5,7%	0,15
Chefe sem instrução	5,5%	6,3%	0,8%	0,83
Chefe com estudo fundamental completo	16,4%	11,6%	-4,8%	0,35
Chefe com nível médio completo	11,0%	11,6%	0,6%	0,89
Chefe com nível superior	0,0%	0,0%	0,0%	
Presença de cônjuge	63,0%	64,3%	1,3%	0,86
Número de filhos	0,86	1,07	20,8	0,14
Chefe tem carteira de identidade	100,0%	99,1%	-0,9%	0,42
Chefe tem CPF	100,0%	100,0%	0,0%	

(continua)

(continuação)

Pré-programa (2002)	Grupo de tratamento	Grupo de comparação	Diferença	Valor-p
Chefe ocupado	47,9%	53,6%	5,6%	0,46
Chefe desocupado	0,0%	1,8%	1,8%	0,25
Chefe inativo	52,1%	44,6%	-7,4%	0,33
Chefe ocupado no setor formal	37,1%	35,0%	-2,1%	0,84
Rendimento do trabalho do chefe	235,99	303,18	67,19	0,25
Características do chefe e cônjuge				
Cônjuge tem carteira de identidade	97,8%	95,8%	-2,0%	0,56
Cônjuge tem CPF	97,8%	95,8%	-2,0%	0,56
Idade do cônjuge	54,83	48,96	-5,87	0,01**
Cônjuge ocupado	37,0%	35,7%	-1,3%	0,86
Cônjuge desocupado	19,2%	27,7%	8,5%	0,19
Cônjuge inativo	43,8%	36,6%	-7,2%	0,33
Rendimento				
Rendimento domiciliar	1.060,85	1.119,18	58,33	0,63
Rendimento domiciliar <i>per capita</i>	395,13	373,58	-21,55	0,61

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

Nota: ** significativo a 5%.

4 METODOLOGIA PARA AVALIAÇÃO DO IMPACTO DO PROGRAMA DE REGULARIZAÇÃO FUNDIÁRIA

4.1 Variável de interesse: renda domiciliar *per capita*

Para avaliar o impacto do Programa de Regularização Fundiária no bem-estar das famílias da Quinta do Caju, utilizamos a renda domiciliar *per capita* como variável de resultado e *proxy* do bem-estar.

Dessa forma, primeiramente, focamos no rendimento total de todos os moradores do domicílio com 10 anos ou mais de idade.²⁴ Como rendimento total, consideramos, além da remuneração bruta da ocupação principal e de outras ocupações, as aposentadorias e/ou pensões, o seguro-desemprego, a renda proveniente de programas sociais, o rendimento de aluguéis e arrendamentos, e as doações ou ajuda financeira. É importante considerar também o número de moradores na análise, pois o grau de bem-estar do domicílio se altera de acordo com o número de seus membros. Utilizamos, portanto, o rendimento domiciliar *per capita*, que consiste na divisão do rendimento domiciliar pelo número de moradores no domicílio.

24. Vale notar que excluímos do rendimento domiciliar os moradores que eram pensionistas, empregados domésticos e parentes do empregado doméstico.

Foi necessário fazer outra restrição à amostra do painel. Como estamos interessados na renda domiciliar *per capita* e com este valor em logaritmo, o trabalho restringiu-se aos domicílios com renda domiciliar maior que 0, o que corresponde a 98,4% dos domicílios particulares permanentes próprios. Portanto, com essa nova seleção, o painel é composto de informações de 182 domicílios particulares permanentes próprios nos anos de 2002 e 2008.

A renda média dos domicílios da Quinta do Caju era de R\$ 1.096,65 em 2002, e R\$ 1.756,38 em 2008, ambos a preços de agosto de 2008. A renda domiciliar *per capita* média era de R\$ 382,21 em 2002, e R\$ 699,07 em 2008. Cabe ressaltar que os domicílios tinham em média 3,1% e 2,7% moradores em 2002 e 2008, respectivamente.

A tabela 4 apresenta a decomposição da renda domiciliar da Quinta do Caju em 2002 e 2008. Em especial, constata-se que a maior parte do rendimento domiciliar da Quinta do Caju é referente à renda do trabalho, apesar de este percentual relativo ter caído em 2008 – queda de 5 pontos percentuais (p.p.). No entanto, a aposentadoria e a pensão têm apresentado um papel importante e crescente no rendimento domiciliar da Quinta do Caju. Destaca-se também o aumento da participação dos rendimentos advindos de programas sociais na renda domiciliar no período analisado. As outras fontes de rendimento não são significativas na composição do rendimento domiciliar da Quinta do Caju.

TABELA 4

Quinta do Caju: decomposição da renda domiciliar – 2002 e 2008

Renda domiciliar média	2002		2008	
	R\$ ¹	%	R\$ ¹	%
Trabalho	666,01	60,8	978,29	55,7
Aposentadoria	263,13	24,0	472,74	26,9
Pensão	120,04	11,0	208,62	11,9
Aluguel	11,45	1,0	8,11	0,5
Seguro-desemprego	1,31	0,1	5,62	0,3
Outros trabalhos	5,29	0,5	31,46	1,8
Ajuda	17,44	1,6	29,66	1,7
Doação	2,18	0,2	1,35	0,1
Outras fontes	6,03	0,6	9,52	0,5
Programas sociais	3,30	0,3	11,02	0,6
Total	1.096,16	100	1.756,38	100

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

Nota: ¹ Valores a preços de agosto de 2008.

A primeira análise a ser considerada, em relação ao efeito do Programa de Regularização Fundiária no rendimento domiciliar da Quinta do Caju, é observar as diferenças das médias do rendimento domiciliar e do rendimento domiciliar *per capita*, entre os grupos de tratamento e de comparação, antes e depois da implementação do programa.

Vemos, primeiramente, de acordo com a tabela 5, que o rendimento domiciliar médio do grupo de comparação era 5,5% maior que o do grupo de tratamento em 2002. No entanto, em 2008, o rendimento domiciliar médio do tratamento passou a ser 9% maior em relação ao grupo de comparação.

Além disso, quanto ao rendimento domiciliar *per capita* médio, observamos, de acordo com a tabela 6, que os valores do grupo de tratamento, em 2002 e 2008, eram 5,8% e 30,8% maiores em relação aos do grupo de comparação, respectivamente. Há uma diferença crescente do rendimento domiciliar *per capita* médio entre os grupos de tratamento e de comparação. Para saber se essa diferença é significativa, é necessário aplicar o teste das diferenças das médias.

Primeiramente, a tabela 5 apresenta o teste das diferenças das médias para o rendimento domiciliar, entre os grupos de tratamento e de comparação, em 2002 e 2008, para as diferentes parcelas do rendimento domiciliar. A única diferença estatisticamente significativa entre os dois grupos é em relação ao rendimento da aposentadoria, tanto para 2002 quanto para 2008.

A tabela 6 apresenta o teste das diferenças das médias para o rendimento domiciliar total e o rendimento domiciliar *per capita*, entre os grupos de tratamento e de comparação, em 2002 e 2008. Observamos que as diferenças das médias do rendimento domiciliar *per capita* entre os grupos de comparação e de tratamento são negativas nos dois anos, mas apenas estatisticamente significativas para o ano de 2008. Contudo, as diferenças das médias do rendimento domiciliar total entre os grupos não são estatisticamente significativas em nenhum dos dois anos.

Os resultados da tabela 6 apontam para a possibilidade de impacto positivo do Programa de Regularização Fundiária na renda domiciliar *per capita* da Quinta do Caju, uma vez que a diferença da média do rendimento domiciliar *per capita* entre o grupo de tratamento e o de comparação passa a ser estatisticamente significativa no ano de 2008. Em todo caso, para confirmar esse resultado, é necessário investigar melhor a relação de causalidade entre o título de propriedade e a renda domiciliar *per capita* a partir da aplicação de alguns métodos econométricos.

TABELA 5

Quinta do Caju: teste das diferenças para o rendimento domiciliar entre os grupos de tratamento e de comparação – 2002 e 2008

Renda domiciliar média	2002				2008			
	Tratamento	Controle	Diferença	Valor-p	Tratamento	Controle	Diferença	Valor-p
Trabalho	580,83	721,52	140,69	0,19	916,37	1.018,64	102,27	0,55
Aposentadoria	354,14	203,81	-150,33	0,01	594,52	393,36	-201,16	0,03
Pensão	90,80	139,10	48,30	0,23	228,82	195,45	-33,38	0,60
Aluguel	13,81	9,90	-3,91	0,75	10,96	6,25	-4,71	0,60
Seguro-desemprego	0	2,16	2,16	0,42	6,03	5,36	-0,67	0,94
Outros trabalhos	3,31	6,57	3,26	0,41	27,40	34,11	6,71	0,84
Ajuda	11,05	21,61	10,56	0,65	31,23	28,64	-2,59	0,89
Doação	0	3,60	3,60	0,42	1,37	1,34	-0,03	0,99
Outras fontes	6,91	5,46	-1,44	0,81	19,25	3,19	-16,06	0,18
Programas sociais	0	5,45	5,45	0,33	12,97	9,74	-3,23	0,74
Total	1.060,85	1.119,18	58,33		1.848,92	1.696,07	-152,85	
Observações	73	112			73	112		

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

Nota: Valores em reais a preços de agosto de 2008.

TABELA 6

Quinta do Caju: teste das diferenças de rendimento domiciliar *per capita* entre os grupos de tratamento e de comparação – 2002 e 2008

Rendimento	2002				2008			
	Grupo de tratamento	Grupo de comparação	Diferença	Valor-p	Grupo de tratamento	Grupo de comparação	Diferença	Valor-p
Rendimento domiciliar	1.060,85	1.119,18	58,33	0,63	1.848,92	1.696,07	-152,08	0,46
Rendimento domiciliar <i>per capita</i>	395,13	373,58	-21,55	0,61	808,76	618,14	-193,43	0,02
Observações	73	112			73	112		

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

Nota: Valores em reais a preços de agosto de 2008.

4.2 Estratégia de identificação

Buscou-se, neste trabalho, medir o impacto médio do Programa de Regularização Fundiária sobre aqueles que foram beneficiados pelo programa, ou seja, $E[Y_{i1} - Y_{i0} | D = 1]$, onde Y_{i1} é o resultado de interesse (rendimento domiciliar *per capita*) para o domicílio i após a implementação do programa, Y_{i0} é o resultado

do mesmo domicílio caso não fosse beneficiado pelo programa e D é uma variável indicadora que é igual a 1 quando o domicílio pertence ao grupo que foi beneficiado pelo programa.

O problema empírico de avaliação de programas sociais é a ausência de dados contrafactuais, uma vez que não é possível observar $E[Y_{i0} | D = 1]$, isto é, não é possível avaliar o resultado de interesse do domicílio que foi beneficiado pelo programa caso não tivesse sido beneficiado. Só é possível averiguar o resultado dos domicílios que participaram do programa, $E[Y_{i1} | D = 1]$, e o resultado dos domicílios que não foram beneficiados, $E[Y_{i0} | D = 0]$. Se pudéssemos observar o rendimento domiciliar *per capita* do grupo de tratamento, caso ele não fosse coberto pelo programa, a avaliação do Programa de Regularização Fundiária seria direta. Dessa forma, o problema empírico concentra-se na estimação do contrafactual: $E[Y_{i0} | D = 0]$ (WOOLDRIDGE, 2000, 2002).

Além disso, vale destacar que temos informação sobre os dois grupos: os que aderiram ao programa $D = 1$ e os que não aderiram ao programa $D = 0$, antes da implementação do programa.

Os três métodos econométricos utilizados neste trabalho – Mínimos Quadrados Ordinários (MQO) agrupado, diferenças em diferenças e *propensity score matching* – baseiam-se na hipótese de que, uma vez que se controle por um vetor de características X , que envolva os fatores, fora do programa, que possam influenciar a renda domiciliar *per capita*, o grupo de comparação pode ser usado como contrafactual.

Todos os métodos serão aplicados nos três painéis: *i*) domicílios próprios; *ii*) domicílios próprios e mesma família; e *iii*) domicílios próprios e mesmo chefe. Esta estratégia será importante para observar a robustez dos resultados.

Por fim, procura-se analisar a robustez dos resultados estimados pelo método de *propensity score matching*. Em especial, analisaremos a sensibilidade dos resultados do modelo estimado por *propensity score matching*, utilizando o Método de Limites de Rosenbaum.

4.2.1 Diferenças em diferenças

O método de diferenças em diferenças fornece o efeito médio do título de propriedade sobre a renda domiciliar *per capita* dos domicílios que participaram do programa, comparando a média da renda domiciliar *per capita* do grupo de tratamento antes e depois do programa com os mesmos resultados dos domicílios do grupo de comparação. O efeito é obtido pela subtração de duas diferenças, isto é:

$$\alpha_{DD} = (\bar{Y}_1^T - \bar{Y}_0^T) - (\bar{Y}_1^C - \bar{Y}_0^C) \quad (1)$$

A primeira diferença elimina os efeitos da heterogeneidade entre os domicílios do grupo de tratamento (efeitos fixos). A segunda diferença só reflete, por hipótese, a influência dos determinantes não observáveis que variaram nesse período. Supondo que os efeitos temporais são iguais para ambos os grupos e que não há mudanças de composição nos grupos, a diferença das diferenças resulta no efeito do título sobre o rendimento domiciliar *per capita* dos domicílios da Quinta do Caju que participaram do Programa de Regularização Fundiária.

A estimação do efeito do programa, considerando as variáveis de controle, pode ser obtida a partir da seguinte regressão:

$$Y_{it} = \alpha D_{gt} + \beta X_{it} + a_i + \varepsilon_{it} \quad (2)$$

onde Y_{it} é uma variável que indica a variável de interesse (rendimento domiciliar *per capita*); D_{gt} uma variável indicadora igual à unidade quando o domicílio pertence ao grupo de tratamento ($g = 1$ e $t = 1$); X_{it} é um vetor de variáveis de controle introduzido para isolar a influência de diferenças entre as características observáveis dos domicílios; a_i é um vetor com as características não observáveis constantes no tempo, e ε_{it} é o termo de erro relacionado às características não observáveis dos domicílios (WOOLDRIDGE, 2000, 2002).

No caso do método das diferenças em diferenças, consideram-se, para o vetor X_{it} , as variáveis observáveis do domicílio que podem influenciar a renda, mas que mudam no tempo. Não consideramos as características do chefe, como sexo e escolaridade,²⁵ pois estas características, em geral, não variam no tempo e, portanto, quando fazemos as diferenças, elas são eliminadas. Dessa forma, preferimos não incluir nenhum controle no modelo.

Em uma regressão linear, o efeito estimado do programa é dado pelo coeficiente α , uma vez que:

$$E(Y_{it} | t = 1, g = 0) - E(Y_{it} | t = 0, g = 0) = \beta E(X_{i1}^C - X_{i0}^C) \quad (3)$$

$$E(Y_{it} | t = 1, g = 1) - E(Y_{it} | t = 0, g = 1) = \alpha + \beta E(X_{i1}^T - X_{i0}^T) \quad (4)$$

25. A escolaridade dos chefes dos domicílios da Quinta do Caju pode ser considerada constante no período de análise, pois os chefes são pessoas com idade média de 60,7 anos.

E sob a hipótese de que

$$E(X_{i1}^T - X_{i0}^T) = E(X_{i1}^C - X_{i0}^C) \quad (5)$$

Temos que:

$$\begin{aligned} & [E(Y_{it} | t=1, g=1) - E(Y_{it} | t=0, g=1)] - \\ & - [E(Y_{it} | t=1, g=0) - E(Y_{it} | t=0, g=0)] = \alpha \end{aligned} \quad (6)$$

Dessa forma, o impacto médio do Programa de Regularização Fundiária é α , ou seja, é a diferença média da renda domiciliar *per capita*, condicional em X , dos domicílios tratados, que participaram do programa, em relação aos não tratados.

A principal limitação do método de diferenças em diferenças diz respeito à construção do grupo de comparação. O método de diferenças em diferenças requer que os grupos de comparação e de tratamento sejam “comparáveis”, ou seja, que as diferenças entre os domicílios dos dois grupos sejam basicamente diferenças de características observáveis controladas de forma linear pela introdução dos controles (HECKMAN; ROBB, 1985).

A precisão e o grau de identificação do impacto do programa sobre os domicílios beneficiados dependem do tamanho do viés de seleção envolvido na avaliação. Essa magnitude, por sua vez, está vinculada ao mecanismo de seleção do grupo de controle ou às técnicas de avaliação utilizadas nessa tarefa (SILVA; RESENDE; SILVEIRA NETO, 2009, p. 94).

Apesar de os grupos de tratamento e de comparação não terem sido construídos de forma aleatória, ou seja, os próprios moradores se autosselecionaram para participar do programa, observamos pelo teste das diferenças de médias, na seção anterior, que as diferenças das características observáveis entre esses dois grupos, antes da implementação do programa, não eram estatisticamente significativas.²⁶ Dessa forma, tomando como base as características observáveis, podemos considerar que os grupos de tratamento e de comparação são de certa forma comparáveis.

Apesar disso, como o método de diferenças em diferenças baseia-se fundamentalmente no uso de um grupo de domicílios não afetados pelo programa que seja comparável ao grupo de tratamento, esse método pode não ser o mais adequado para a análise.

26. No entanto, vale apontar que as diferenças da idade do chefe e do cônjuge, em 2002, eram estatisticamente significativas entre os grupos de tratamento e de comparação.

4.2.2 Propensity score matching

Considerando que a hipótese $E(X_{i1}^T - X_{i0}^T) = E(X_{i1}^C - X_{i0}^C)$ não seja válida, pois os grupos de tratamento e de comparação não são tão parecidos em relação às características observáveis, uma vez que eles não foram selecionados aleatoriamente, o resultado estimado por meio do método das diferenças em diferenças apresentará um viés de autosseleção. Uma forma de contornar esse viés é estimar o efeito do programa por meio do método de *propensity score matching*.

O *propensity score matching* contrasta a renda domiciliar *per capita* de um domicílio que participou do Programa de Regularização Fundiária com a renda domiciliar *per capita* de um domicílio (ou um grupo de domicílios) com características similares do grupo de comparação. Para tanto, esse método calcula a probabilidade de os domicílios da amostra participarem do programa, o *propensity score*, baseado em um vetor X de características observáveis.

Portanto, o *propensity score*, $P(X)$, é definido como a probabilidade condicional de um indivíduo receber o tratamento dadas suas características observáveis X . Isto é,

$$P(X) = \text{probabilidade}(D = 1 | X).^{27} \quad (7)$$

Se o tratamento e os resultados esperados são independentes condicionais às variáveis de pré-tratamento, estes também serão independentes condicionais à probabilidade de receber o tratamento tomando como base as características observáveis, isto é, condicionais ao *propensity score*:

$$[Y_{i0}, Y_{i1} \perp\!\!\!\perp D | P(X)] \quad (8)$$

A utilização do *propensity score* baseia-se em duas hipóteses-chave. A hipótese essencial dessa abordagem é que o resultado da variável de interesse dos não tratados é independente do *status* de participação, D , uma vez que se controle pelas variáveis observáveis, X , por meio da função $P(X)$:

$$E(Y_{i0} | P(X), D = 1) = E(Y_{i0} | P(X), D = 0) \quad (9)$$

Ou seja, dado $P(X)$, o resultado dos não tratados é igual ao hipotético resultado dos domicílios tratados caso não tivessem passado pelo tratamento (HECKMAN; ICHIMURA; TODD, 1998).

27. D é uma variável *dummy* que possui valor igual a 1, se o indivíduo participa do programa, e valor 0, caso ele não participe.

A segunda hipótese refere-se à existência de um suporte comum. Esta condição requer que existam unidades de ambos os grupos, tratamento e controle, para cada característica X para a qual se deseja comparar: $0 < P(X) < 1$ (HECKMAN; ICHIMURA; TODD, 1998).

O objetivo do *matching* é selecionar fatores observáveis suficientes, de forma que dois domicílios com os mesmos valores desses fatores não terão respostas sistematicamente distintas em relação à introdução do programa.

O modelo *logit* (ou *probit*), pelo qual o *propensity score* é calculado, deve incluir variáveis preditoras que influenciam a participação no programa e os resultados de interesse. Consideraremos como vetor X_{it} as variáveis referentes às características do chefe do domicílio em 2002, antes da implementação do programa. Em particular, utilizaremos as seguintes variáveis: idade do chefe; *dummy* que indica o sexo do chefe (assume valor 1 para as mulheres); escolaridade do chefe (anos de estudo); três *dummies* para a ocupação do chefe (inativo, ocupado formal, desocupado), tomando como base o chefe ocupado no setor informal; *dummy* para participação social²⁸ do chefe; *dummy* para presença de cônjuge no domicílio; e variável que indica o número de filhos que moram no domicílio.

O *propensity score* estimado fornece a probabilidade predita de participação no programa de um determinado domicílio e por meio dos seus valores realiza-se o *matching* entre os grupos de tratamento e de comparação. A inclusão de variáveis com as características do chefe tem por finalidade não só o cálculo da probabilidade de participação, mas também o *matching* dos domicílios em termos das características observáveis.

Uma vez calculado o *propensity score*, mede-se a diferença da renda domiciliar *per capita* de um domicílio do grupo de tratamento com a renda de um domicílio (ou a média da renda de um grupo de domicílios) do grupo de comparação com *propensity score* semelhante. Assim, usando o *propensity score*, o grupo de comparação para cada indivíduo é construído baseado num critério predefinido de proximidade. Existem diversos estimadores de *propensity score matching*, e a diferença crucial é a definição de proximidade, ou melhor, o tipo de peso utilizado. O efeito total da introdução do programa é a média dos efeitos individuais dos domicílios beneficiados pelo programa.

Em suma, a estimação do efeito do tratamento sobre os tratados usando *propensity score* necessita basicamente da estimação de $P(X)$, calculada por meio de um *logit*, e da escolha de pesos apropriados, que inclui uma série de alternativas como pesos de *kernel*, pesos baseados na vizinhança mais próxima ou dentro de um raio especificado.

28. Consideramos a participação social do chefe em atividades em instituições como, por exemplo, partidos políticos, sindicatos, associações de moradores, cooperativas, igrejas, Organizações Não Governamentais (ONGs), entre outras.

Em geral, na literatura de avaliação, o parâmetro que recebe maior atenção é o que representa o efeito do tratamento ou efeito médio do tratamento nos tratados – *average treatment effect on the treated* (ATT).

$$ATT = E[Y_{1i} - Y_{0i} | D = 1, P(X)] = E[Y_{1i} | D = 1, P(X)] - E[Y_{0i} | D = 1, P(X)] \quad (10)$$

Outro parâmetro de interesse é o efeito médio do tratamento – *average treatment effect* (ATE) – que é definido como igual a $E[Y(1) - Y(0)]$. Dessa forma, enquanto o ATT é o efeito do tratamento apenas sobre os tratados, o ATE é o efeito do tratamento sobre toda a amostra.

5 RESULTADOS

Nesta seção, apresentaremos os resultados da estimação do impacto do Programa de Regularização Fundiária no rendimento domiciliar *per capita* da Quinta do Caju, por meio dos métodos das diferenças em diferenças e do *propensity score matching*.

5.1 Resultados por diferenças em diferenças

Os resultados da estimação da equação das diferenças em diferenças para o rendimento domiciliar *per capita* da Quinta do Caju estão apresentados na tabela 7. A primeira coluna da tabela 7 apresenta os resultados para o painel em que são considerados todos os domicílios próprios da Quinta do Caju. A segunda e a terceira colunas restringem a amostra para os domicílios com a mesma família e o mesmo chefe nos dois períodos da análise (antes e depois da implementação do programa), respectivamente.

TABELA 7

Quinta do Caju: equação de diferenças em diferenças do rendimento domiciliar *per capita* – 2002 e 2008

Var lnRDPC	Mesmo domicílio	Mesma família	Mesmo chefe
Participação no programa	0,123 (1,21)	0,115 (1,07)	0,139 (1,15)
Constante	0,586 (9,28)**	0,601 (9,04)**	0,587 (7,76)**
Observações	182	164	136
R^2	0,01	0,01	0,01

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

Notas: Estatística-t robusta entre parênteses.

** significativo a 1%.

Observa-se que o coeficiente da variável binária que indica a participação no programa é positivo, mas não estatisticamente significativo, em nenhum dos painéis analisados.

Como os grupos de controle e comparação não foram selecionados de forma experimental ou aleatória, os resultados com o método das diferenças em diferenças sugerem o uso do método alternativo *propensity score matching*, em que a comparação do rendimento domiciliar *per capita* é feita entre domicílios mais similares. Os resultados com o método de *propensity score matching* será apresentado a seguir.

5.2 Resultados por *propensity score matching*

Primeiramente, para medir o impacto do Programa de Regularização Fundiária no rendimento domiciliar *per capita* dos domicílios da Quinta do Caju beneficiados pelo programa, por meio do método do *propensity score matching*, foi estimado o *propensity score* com uma regressão *logit*, com base no vetor X_{it} , com as variáveis referentes às características do chefe em 2002 (antes da implementação do programa).²⁹

Após calcular o *propensity score*, estimamos a diferença do logaritmo da renda domiciliar *per capita* de um domicílio do grupo de tratamento com um domicílio (ou um grupo de domicílios) do grupo de comparação com *propensity score* semelhante, utilizando quatro estimadores diferentes de *propensity score matching*: *Nearest Neighbor* com apenas um vizinho [*NN*(1)], *kernel* gaussiano com janela de 0,06 [*Kernel* (0,06)], *kernel* gaussiano com janela de 0,1 [*Kernel* (0,1)] e *kernel* gaussiano com janela de 1 [*Kernel* (1)].

A tabela 8 apresenta os resultados da regressão *logit*, que são os mesmos para todos os quatro métodos de *propensity score matching*, já que usamos a mesma regressão para calcular o *propensity score*. A primeira coluna da tabela 8, considerando os domicílios próprios da Quinta do Caju, revela que apenas o coeficiente referente à idade do chefe em 2002 é estatisticamente significativo ao nível de 5%. O coeficiente da variável sexo do chefe é estatisticamente significativo ao nível de 10%. Espera-se, para cada um ano a mais na idade do chefe, um aumento de 0,05 no *log-odds* de participar no programa. Além disso, os chefes homens têm probabilidade maior de participar do programa.

Quanto aos resultados do *logit* para os domicílios próprios com a mesma família ou com o mesmo chefe, observa-se que apenas o coeficiente referente à idade do chefe em 2002 é estatisticamente significativo ao nível de 5% para os domicílios com a mesma família. Nesse caso, espera-se, para cada um ano a mais na idade do chefe, um aumento de 0,04 no *log-odds* de participar no programa, ou seja, um aumento menor em relação ao resultado do painel dos domicílios próprios.

29. Em particular, utilizamos as seguintes variáveis: idade do chefe, *dummy* que indica o sexo do chefe (assume valor 1 para as mulheres); escolaridade do chefe (anos de estudo), três *dummies* para a ocupação do chefe (inativo, ocupado formal, desocupado), tomando como base o chefe ocupado no setor informal, *dummy* para participação social do chefe; *dummy* para presença de cônjuge no domicílio, e variável que indica o número de filhos que moram no domicílio.

TABELA 8

Resultados do *logit*: participação no programa

Características do chefe em 2002	Mesmo domicílio	Mesma família	Mesmo chefe
sexo do chefe (mulher)	-1,451 (-1,72)*	-1,401 (-1,38)	-1,274 (-1,21)
idade do chefe	0,046 (2,31)**	0,043 (2,04)**	0,036 (1,59)
escolaridade do chefe	0,795 (1,43)	0,532 (0,92)	0,042 (0,61)
Chefe inativo	-0,022 (-0,05)	0,425 (0,10)	0,469 (0,98)
Chefe ocupado no setor formal	0,319 (0,66)	0,350 (0,70)	0,680 (1,28)
Participação social do chefe	0,119 (0,31)	0,026 (0,07)	-0,090 (-0,21)
Presença de cônjuge	-1,191 (-1,49)	-1,408 (-1,45)	-0,827 (-0,82)
Número de filhos	-0,036 (-0,19)	-0,023 (-0,13)	-0,101 (-0,47)
Constante	-2,549 (-1,73)	-1,992 (-1,27)	-2,115 (-1,20)
Observações	180	162	135
Prob > chi2	0,1728	0,3697	0,4127
Pseudo R ²	0,0476	0,0397	0,0452

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

Notas: Estatística-z entre parênteses.

O grupo de referência para a ocupação do chefe é ocupado no setor informal. A variável binária desocupado foi excluída do modelo por apresentar poucas observações.

* significativo a 10%; ** significativo a 5%.

No resultado do *logit*, verificou-se que a maioria das variáveis independentes não é estatisticamente significativa, mas isso não significa necessariamente que estas variáveis não devam ser incluídas no modelo. Rubin e Thomas (1996) (*apud* CALIENDO; KOPEINIG, 2005) argumentam que uma variável só deve ser excluída da análise se houver um consenso de que esta não é apropriada (não relacionada com o resultado). De outra forma, é aconselhável incluí-la na estimativa do *propensity score* mesmo que não seja estatisticamente significativa. Segundo Caliendo e Kopeinig (2005), a introdução de variáveis não estatisticamente significativas não viesia nem torna inconsistentes as estimativas.

As tabelas 9, 10 e 11 apresentam as estimativas do impacto do programa sobre a variação da renda domiciliar *per capita*³⁰ para os quatro métodos utilizados na estimação

30. $\text{var}(\ln RDPC) = \ln RDPC_{2008} - \ln RDPC_{2002}$.

do *propensity score matching*, e para os três tipos de painéis. A análise do impacto será feita com base em quatro pontos: *i*) a magnitude do valor estimado para o ATT; *ii*) o sinal do ATT; *iii*) a significância estatística do ATT; e *iv*) o valor estimado para o ATE.

TABELA 9

***Propensity score matching* para os domicílios próprios da Quinta do Caju – 2002 e 2008**

VarlnRDPC	<i>NV</i> (1)	<i>Kernel</i> (0,06)	<i>Kernel</i> (0,1)	<i>Kernel</i> (1)
ATT				
Tratamento	0,709	0,709	0,726	0,726
Controle	0,433	0,501	0,542	0,565
Diferença	0,276	0,208	0,184	0,161
Estatística-t	2,88	1,95	1,74	1,54
ATE	0,174	0,195	0,188	0,162
Observações	180	180	180	180

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

TABELA 10

***Propensity score matching* para os domicílios próprios da Quinta do Caju com a mesma família – 2002 e 2008**

VarlnRDPC	<i>NV</i> (1)	<i>Kernel</i> (0,06)	<i>Kernel</i> (0,1)	<i>Kernel</i> (1)
ATT				
Tratamento	0,715	0,715	0,727	0,727
Controle	0,580	0,568	0,585	0,587
Diferença	0,136	0,148	0,142	0,140
Estatística-t	1,36	1,33	1,28	1,28
ATE	0,113	0,151	0,148	0,140
Observações	162	162	162	162

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

TABELA 11

***Propensity score matching* para os domicílios próprios da Quinta do Caju com o mesmo chefe – 2002 e 2008**

VarlnRDPC	<i>NV</i> (1)	<i>Kernel</i> (0,06)	<i>Kernel</i> (0,1)	<i>Kernel</i> (1)
ATT				
Tratamento	0,726	0,726	0,744	0,744
Controle	0,551	0,514	0,542	0,555
Diferença	0,175	0,212	0,202	0,189
Estatística-t	1,52	1,68	1,64	1,56
ATE	0,126	0,189	0,199	0,190
Observações	135	135	135	135

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

As colunas das tabelas referem-se aos diferentes métodos: *NN* (1), *kernel* (0,06), *kernel* (0,1) e *kernel* (1). Além disso, as tabelas reportam os resultados destes métodos para os tratados, o contrafactual (a estimativa dos resultados para os tratados, caso eles não tivessem recebido o tratamento), o ATT e a estatística-t (para a hipótese nula de que não há diferença nos resultados para o grupo de tratamento e o contrafactual), assim como o resultado do ATE.

Observa-se, portanto, que todas as estimativas do ATT são positivas para todos os métodos e em todos os painéis. Primeiramente, utilizando o método *NN* com apenas um vizinho (*NN*(1)) para os domicílios próprios da Quinta do Caju, o ATT estimado é de 0,276. Isto representa um aumento da renda domiciliar *per capita* de aproximadamente 31,8³¹ nos domicílios tratados em relação a um grupo similar, que não recebeu o tratamento. Este resultado é estatisticamente significativo ao nível de 1% (estatística-t de 2,88). Vale apontar que o valor do ATE é de 0,174.

Em relação ao método do vizinho mais próximo, o método de *kernel*, ao estimar o ATT, utiliza mais informações, pois considera a média de todas as observações para criar o contrafactual (usando os pesos de *kernel* – *kernel weights*). No entanto, como nem todas as observações do grupo de comparação são parecidas com aquela observação específica do grupo de tratamento, o método de *kernel* utiliza informações de domicílios mais distantes em termos das características do *logit* usado no cálculo do *propensity score*.

Primeiramente, com o *kernel* com janela de 0,06, observa-se que o ATT estimado é de 0,208, com o logaritmo da renda domiciliar *per capita* aumentando de 0,501 para 0,709. Ou seja, estima-se um aumento da renda domiciliar *per capita* de aproximadamente 23,1 nos domicílios tratados em relação aos não tratados. Este resultado é estatisticamente significativo a 10% (estatística-t de 1,95). Vale notar ainda que o ATE é de 0,195.

Em segundo lugar, com o *kernel* com janela de 0,1, constata-se que o ATT estimado é de 0,184, de uma base de 0,542 para o grupo dos não tratados. Ou seja, estima-se um aumento da renda domiciliar *per capita* de aproximadamente 20,2 nos domicílios tratados em relação aos não tratados. Esse efeito é estatisticamente significativo a 10% (estatística-t de 1,74). Vale ressaltar que o valor do ATE é de 0,188.

Por fim, aumentando a janela do *kernel* para 1, ou seja, dando peso maior para os domicílios ainda menos “parecidos”, o ATT estimado passa a ser de 0,161, de uma base de 0,565. No entanto, o resultado não é estatisticamente significativo em níveis convencionais.

31. Como os coeficientes têm valores elevados, a interpretação em termos de variação percentual não é aproximada. Dessa forma, foram feitos os ajustes necessários, usando $[\exp(\beta) - 1]$.

Constata-se que, quanto maior a janela do *kernel*, menor é o valor do ATT e o resultado passa a ser estatisticamente menos significativo. Além disso, pelo método do vizinho mais próximo (*NN(01)*) para os domicílios próprios da Quinta do Caju, o valor do ATT é maior do que o estimado pelo método de *kernel*, assim como mais significativo estatisticamente.

Considerando os domicílios próprios da Quinta do Caju com a mesma família, constata-se que o ATT é positivo, mas não é estatisticamente significativo em nenhum dos modelos de *propensity score matching* (tabela 10). Esse resultado não significativo pode ser devido à mudança da composição familiar no período.

Considerando os domicílios próprios da Quinta do Caju com o mesmo chefe (tabela 11), observa-se que o ATT é positivo e significativo estatisticamente ao nível de 10% para os métodos de *kernel* com janela de 0,06 e 0,1. Nesse caso, estima-se um aumento da renda domiciliar *per capita* em torno de 21 e 22 nos domicílios tratados em relação aos não tratados.

Conclui-se que todos os métodos utilizados de estimação do ATT apontam para um efeito positivo e estatisticamente significativo do Programa de Regularização Fundiária na variação da renda domiciliar *per capita* dos domicílios próprios da Quinta do Caju que participaram do programa, com exceção do painel para as mesmas famílias.

Na próxima seção, analisaremos se estes resultados estimados por meio do método de *propensity score matching* são robustos.

5.3 Avaliação da robustez dos resultados

Nesta subseção, avaliaremos a robustez dos resultados estimados por meio do método de *propensity score matching*. Para tanto, primeiramente, apresentaremos o método de limites de Rosenbaum (ROSENBAUM, 2002), que avalia o impacto potencial do viés de seleção que surge devido às variáveis não observadas. Em seguida, aplicaremos este método para observar se os resultados do impacto do programa sobre o rendimento domiciliar *per capita* são robustos.

5.3.1 Análise de sensibilidade: o método de limites de Rosenbaum

Os estimadores baseados no *propensity score* não são estimadores consistentes do efeito do tratamento se a participação no programa é endógena, isto é, se variáveis não observadas que afetam o processo de seleção também afetam os resultados (DIPRETE; GANGL, 2004, p. 22). Os fatores não observados (como, por exemplo, a motivação), provavelmente, se correlacionam com a participação no programa e com a variável de resultado. Portanto, a omissão dessas características individuais pode causar um viés nos resultados do efeito médio do tratamento sobre o tratado, estimado sob a hipótese de exogeneidade (RESENDE; OLIVEIRA, 2008).

Como não é possível estimar a magnitude do viés de seleção em dados não experimentais, uma importante ferramenta desenvolvida por Rosenbaum (2002) pode ser empregada para avaliar esse viés na estimação dos efeitos causais. Dessa forma, a análise de limites (*bounds analysis*) avalia o impacto potencial do viés de seleção que surge devido às variáveis não observadas, ao determinar quão forte deve ser a influência de uma variável omitida sobre a seleção na participação, de modo a prejudicar as conclusões a respeito dos efeitos causais, obtidos com as análises de *matching* (CALIENDO; KOPEINIG, 2005, p. 19).

Em suma, a análise de sensibilidade pode ser utilizada para testar a robustez dos resultados na presença de viés devido a uma covariável omitida. O objetivo da análise de sensibilidade é avaliar como uma variável não observável que afeta tanto a participação no Programa de Regularização Fundiária como os resultados (variação no logaritmo do rendimento domiciliar *per capita*) pode alterar as conclusões em relação ao efeito do tratamento.

Considerando-se que a probabilidade de participação de um domicílio i no programa seja dada por (CALIENDO; KOPEINIG, 2005, p. 20; RESENDE; OLIVEIRA, 2008, p. 245-246):

$$P(x_i) = P[(D = 1)|x_i] = F(\beta x_i + \gamma u_i) \quad (11)$$

onde x_i são as características observáveis do chefe do domicílio, D é igual a 1 se o domicílio recebe o tratamento e 0 caso não receba; u_i corresponde à variável não observada; e γ representa o efeito de u_i sobre a decisão de participação no programa. Se não existir viés de seleção, então γ será igual a 0 e a probabilidade de participação será exclusivamente determinada pelas características observáveis. Entretanto, na presença de viés de seleção, dois domicílios com as mesmas covariáveis observadas, x , terão chances diferentes de participar no programa.

Assumindo que dois domicílios sejam pareados, i e j , e que F tenha uma distribuição logística, a *odds* de os domicílios participarem do programa é dada por:

$$\frac{P(x_i)}{[1 - P(x_i)]} = \frac{P(x_j)}{[1 - P(x_j)]}$$

E a *odds ratio* é dada por:

$$\frac{\frac{P(x_i)}{[1 - P(x_i)]}}{\frac{P(x_j)}{[1 - P(x_j)]}} = \frac{P(x_i)[1 - P(x_j)]}{P(x_j)[1 - P(x_i)]} = \frac{\exp(\beta x_j + \gamma u_j)}{\exp(\beta x_i + \gamma u_i)} = \exp[\gamma(u_i - u_j)] \quad (12)$$

Se os domicílios possuírem as mesmas características observáveis, então o vetor x se cancela. Desta forma, se não houver diferenças nas variáveis não observadas ($u_i = u_j$) e se estas variáveis não influenciarem a probabilidade de participação ($\gamma = 0$), a *odds ratio* será igual a 1, implicando a não existência de viés de seleção. Se a *odds* de participação diferir – isto é, se a *odds ratio* for diferente de 1 – só pode ser devido à presença de não observáveis. A análise de sensibilidade avalia o quanto do efeito do programa é alterado pela mudança nos valores de γ e de $u_i - u_j$.

Na prática, devem-se examinar os limites da *odds ratio* de participação. Rosenbaum (2002) mostra que a equação (13) implica os seguintes limites para a *odds ratio*:

$$\frac{1}{e^\gamma} \leq \frac{P(x_i)[1 - P(x_j)]}{P(x_j)[1 - P(x_i)]} \leq e^\gamma \quad (13)$$

Os domicílios pareados possuem a mesma probabilidade de participação apenas se $e^\gamma = 1$. Entretanto, se $e^\gamma = 2$, então, domicílios aparentemente similares, em termos de x , diferirão em suas probabilidades de participar do programa por um fator de até 2. Assim, segundo Rosenbaum (2002), e^γ (ou gamma = $\Gamma = e^\gamma$) seria uma medida do grau de ruptura a partir de um estudo livre de viés de seleção.

5.3.2 Análise de sensibilidade: resultados

Segundo Diprete e Gangl (2004), a metodologia empregada inicia-se com a estimativa do ATT utilizando métodos de *propensity score*, baseada na hipótese de independência condicional. Em seguida, postula-se a existência de uma variável omitida Z , que afeta a probabilidade de participação no programa (isto é, a probabilidade de $D = 1$), condicional em X . Conforme o impacto potencial de Z sobre D (expresso em termos dos limites das *odds ratio*) torna-se mais forte, o intervalo de confiança dos efeitos estimados amplia-se, e o nível de significância do teste da hipótese nula, de que não há efeito de D sobre Y (variável de interesse), aumenta, ou seja, o valor-p eleva-se. Em outras palavras, quanto maior for o valor do gamma ($\Gamma = e^\gamma$), maior será o valor-p e, conseqüentemente, menor será a significância

estatística do efeito do programa sobre a variação da renda domiciliar *per capita* dos tratados.

O resultado da tabela 12 mostra a robustez dos diferentes métodos de *propensity score matching*, ao viés de seleção, para a variável de resultado (variação do rendimento domiciliar *per capita*).³²

Vale ressaltar que aplicamos este método de análise de sensibilidade apenas para os resultados estimados com o painel de domicílios próprios, pois foi o que apresentou estimativas mais estatisticamente significativas para o ATT.

TABELA 12

Quinta do Caju: análise de sensibilidade (*Rosenbaum bounds*) para a variação no rendimento domiciliar *per capita* – 2002 e 2008

<i>NN</i> (1)		<i>Kernel</i> (0,06)		<i>Kernel</i> (0,1)		<i>Kernel</i> (1)	
Γ	p-crítico	Γ	p-crítico	Γ	p-crítico	Γ	p-crítico
1,00	0,019	1,00	0,022	1,00	0,043	1,00	0,077
1,05	0,029	1,05	0,034	1,05	0,062	1,05	0,106
1,10	0,042	1,10	0,048	1,10	0,085	1,10	0,140
1,15	0,059	1,15	0,067	1,15	0,113	1,15	0,179
1,20	0,079	1,20	0,089	1,20	0,144	1,20	0,221
1,25	0,103	1,25	0,116	1,25	0,180	1,25	0,267
1,30	0,130	1,30	0,145	1,30	0,219	1,30	0,315

Fontes: PSCBRC, elaborada em 2002 pelo IETS/Firjan/Science, e Pesquisa Socioeconômica da Comunidade da Quinta do Caju, elaborada em 2008, com financiamento da FAPERJ.

Nota: A segunda coluna de cada método contém os valores de p+ para os limites de Rosenbaum. Se $\Gamma = e^y = 1$ implica a não existência de viés devido a não observações.

Considerando o nível de significância de no máximo 10%, as estimativas de p-crítico observadas na tabela 12 indicam que o valor do gamma crítico está entre 1,20 e 1,25 para os métodos *NN*(1) e *Kernel* (0,06), entre 1,10 e 1,15 para o *Kernel* (0,1) e entre 1 e 1,05 para o *Kernel* (1).³³ Dessa forma, a estimação do efeito do tratamento nos tratados parece ser robusta a uma possível presença de viés de seleção, pois considerando que quanto maior o valor do nível crítico Γ , mais robusta é a estimativa de ATT.

32. Considerando $e^y = \Gamma$ e adotando a estratégia de Diprete e Gangl (2004, p. 22): *At each Γ we calculate a hypothetical significance level "p-critical", which represents the bound on the significance level of the treatment effect in the case of endogenous self-selection into treatment status. By comparing the Rosenbaum bounds on treatment effects at different levels of Γ we can assess the strength such unmeasured influences would require in order that the estimated treatment effects from propensity score matching would have arisen purely through selection effects.*

33. Ao contrário do encontrado nos artigos de Diprete e Gangl (2004, p. 55) e Resende e Oliveira (2008, p. 260), o gamma crítico neste trabalho foi determinado mais em função de um valor limite para o nível de significância (10%) do ATT do que a partir de uma mudança abrupta nos valores do nível de significância para um determinado acréscimo no valor de gamma.

Portanto, conclui-se que as estimativas pelo método de *Nearest Neighbor* com apenas um vizinho ($NN(1)$) e *kernel* com janela 0,06 parecem ser as mais robustas para o efeito do programa no rendimento domiciliar *per capita* dos tratados.

Além disso, Diprete e Gangl (2004) destacam que estes resultados são considerados como os piores cenários. Dessa forma, o valor do γ crítico de próximo de 1,05 para o método de *Kernel* (1), por exemplo, não sugere necessariamente que não exista um efeito positivo sobre a renda domiciliar *per capita*. Segundo os autores, isso tem como consequência somente que o intervalo de confiança para o efeito do tratamento incluirá 0 se a presença de viés de seleção implicar que a *odds* de participação seja diferente entre os grupos de tratamento e de controle por um fator de até 1,05. Quando uma variável omitida possui forte influência sobre a participação no programa, mas apenas uma pequena influência sobre a variável de resultado, o intervalo de confiança não contém 0. Apesar de a análise de sensibilidade de Rosenbaum apresentar os resultados para o pior cenário, ele demonstra quão grande deve ser a influência de uma variável não observada de modo a questionar as conclusões obtidas por meio de métodos de *propensity score matching*.

6 CONCLUSÃO

Considerando a relevância dos direitos de propriedade para o bem-estar, buscamos verificar os impactos do Programa de Regularização Fundiária no bem-estar das famílias da Quinta do Caju, no Rio de Janeiro. Este programa foi pioneiro na cidade e hoje a Quinta do Caju é a comunidade de baixa renda com maior número de domicílios beneficiados pelo título de propriedade no Brasil.

Em especial, o trabalho estimou, no período entre 2002 e 2008, o impacto do Programa de Regularização Fundiária na renda domiciliar *per capita* dos domicílios próprios da Quinta do Caju que foram beneficiados pelo programa.

A base de dados utilizada foi um painel referente aos mesmos domicílios da Quinta do Caju em 2002 (antes do programa) e em 2008 (depois do programa). O fato da utilização de um painel com os mesmos domicílios antes e depois do programa já é por si só uma contribuição para a literatura de avaliação de impacto dos programas de regularização fundiária no Brasil, que tem se baseado até agora em dados em *cross-section*.

Para essa avaliação, aplicamos dois métodos econométricos diferentes (diferenças em diferenças e *propensity score matching*) em três subconjuntos de uma base de dados em painel dos domicílios próprios da Quinta do Caju: mesmos domicílios, mesma família e mesmo chefe nos dois anos.

Os resultados com o método de diferenças em diferenças não foram estatisticamente significativos, sugerindo que o direito de propriedade mais bem definido não tem impacto sobre a renda domiciliar *per capita* na Quinta do Caju.

No entanto, como os grupos de tratamento e comparação não foram selecionados de forma aleatória, o método de *propensity score matching* pode ser considerado mais adequado, já que permite parear os domicílios do grupo de tratamento e de controle por meio das características observáveis, reduzindo assim o viés de autoseleção. Com este método, os resultados revelam que o efeito médio do tratamento sobre os tratados (ATT) é sempre positivo. Entretanto, dependendo da especificação do modelo e do algoritmo de *matching* adotado, o nível de significância estatística do ATT varia.

Os resultados estimados do ATT indicam um aumento em torno de 20% a 32% da renda domiciliar *per capita* dos domicílios próprios tratados em relação aos não tratados da Quinta do Caju. Esses resultados são robustos quanto à possível presença de viés de seleção pelas variáveis não observáveis com um gamma de até 1,20, considerando a análise de sensibilidade de Rosenbaum.

Os resultados apontam, portanto, para um efeito positivo e estatisticamente significativo dos títulos de propriedade sobre o rendimento domiciliar *per capita* na Quinta do Caju.

A existência de tantas favelas na cidade do Rio de Janeiro aponta certa tolerância do governo em relação à informalidade. Os moradores, apesar de afirmarem que o título lhes daria mais segurança quanto à ocupação, não deixam de investir em seus imóveis ou participar do mercado de trabalho. O sistema judiciário brasileiro, ao proteger o bem familiar, também não permite o uso do imóvel como garantia de empréstimo, no caso das famílias pobres em que o imóvel representa o único bem da família. Dessa forma, o acesso ao crédito ainda é restrito para os moradores de comunidades de baixa renda mesmo com a posse do título de propriedade dos seus imóveis.

A presença do Estado ainda é muito pequena nessas comunidades, a ponto de surgir um impacto significativo no bem-estar das famílias. É importante que o Estado defina, atribua e garanta os direitos de propriedade, assim como facilite o acesso a mercados. Isso possibilitará que os moradores se sintam incluídos, fortalecidos e responsáveis, de forma a investirem na comunidade e buscarem maior integração na economia formal, como, por exemplo, registrar o negócio, utilizar a rede de serviços formais (luz, TV a cabo, internet etc.), contribuir para a previdência, pagar o IPTU etc.

Portanto, apesar dos esforços do governo ao implementar o Programa de Regularização Fundiária, sua ação é bastante tímida, principalmente se considerarmos que foi iniciado por uma pequena favela histórica, homogênea e bem estabelecida na cidade do Rio de Janeiro. Segundo a literatura sobre direitos de propriedade e bem-estar, se o objetivo do governo é aumentar o impacto dos programas de regularização fundiária, deveria intensificar o programa e atender favelas mais heterogêneas e com forte informalidade. Dessa forma, as avaliações de impacto

também serão mais precisas em identificar os canais para o aumento do bem-estar das famílias beneficiadas pelo programa.

Por fim, futuros estudos, a partir de bases de dados que permitam captar melhor os possíveis canais que levam ao aumento da renda domiciliar *per capita* depois de programas de regularização fundiária, podem contribuir para a agenda sobre desenvolvimento socioeconômico, especialmente no Rio de Janeiro.

ABSTRACT

This work analyses the relationship between property rights and welfare. Considering the relevance of institutions in the economy, it is supposed that, if property rights are better defined, attributed and guaranteed in slums, the transaction costs will be low and, consequently, the value of the asset and the income of the owner of this asset will be greater. Further, with the property rights established and guaranteed, the low income communities' residents are then included in the formal city and are given access to a whole new variety of markets. This study presents an impact evaluation of the Land Regularization Program implemented by the City Hall of Rio de Janeiro in the low income community of Quinta do Caju. The program has given land ownership titles, formally registered by notary public, for the beneficiary households. Using panel data with socioeconomic information of the same households of Quinta do Caju before and after the program, the effects of the program on the welfare of residents are estimated. Applying propensity score matching method, it is observed that the program has a positive and statistically significant impact on the per capita household income of beneficiaries.

REFERÊNCIAS

- ALSTON, L. J.; LIBECAP, G. D.; SCHNEIDER, R. The determinants and the impact of property rights: land titles on the Brazilian frontier. **Journal of Law, Economics and Organization**, v. 12, n. 1, p. 25-61, 1996.
- ANDRADE, M. I. de T. **Direitos de propriedade e renda pessoal**: um estudo de caso das comunidades do Caju. 2004. Dissertação (Mestrado) – Instituto de Economia/Universidade Federal do Rio de Janeiro, Rio de Janeiro, set. 2004.
- ANGEL, S. *et al.* **Secure tenure in Latin America and Caribbean**: regularization of informal urban settlements in Peru, Mexico and Brazil. Princeton, NJ: Woodrow Wilson School of Public and International Affairs Princeton University, Jan. 2006.
- BANERJEE, A.; GERTLER, P.; GHATAK, M. Empowerment and efficiency: tenancy reform in West Bengal. **Journal of Political Economy**, v. 110, n. 2, p. 239-280, 2002.
- BESLEY, T. Property rights and investment incentives: theory and evidence from Ghana. **The Journal of Political Economy**, v. 103, n. 5, p. 903-937, Oct. 1995.
- CALIENDO, M.; KOPEINIG, S. **Some practical guidance for the implementation of propensity score matching**. IZA, May 2005 (Discussion Paper, n. 1.588).
- CANTUARIAS, F.; DELGADO, M. **Peru's urban land titling program**. Case study from "Reducing poverty, sustaining growth – what works, what doesn't and why? A global exchange for scaling up success. Shanghai: World Bank, 25-27, May 2004.

CARDOSO, A.; ELIAS, P.; PERO, V. **Urban regeneration and spatial discrimination: the case of Rio's favelas.** In: ENCONTRO NACIONAL DE ECONOMIA, 31., 2003, Salvador. **Anais...** Salvador: ANPEC, dez. 2003.

CARTER, M.; OLINTO, P. **Getting institutions right for whom: credit constraints and the impact of property rights on the quantity and composition of investment.** University of Wisconsin-Madison, Agricultural and Applied Economics Department, 2000 (Staff Paper, n. 433).

_____.; WIEBE, K. D.; BLAREL, B. Tenure security for whom? Differential effects of land policy in Kenya. In: BRUCE, J.; MIGOT-ADHOLLA, S. E. (Ed.). **Searching for land tenure security in Africa.** Dubuque, Iowa: Kendall/Hunt Publishing Company, 1994. cap. 7.

DE SOTO, H. **O mistério do capital.** Rio de Janeiro: Record, 2000.

DIPRETE, T.; GANGL, M. Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. **Sociological Methodology**, v. 34, n. 1, p. 271-310, Apr. 2004.

DO, Q. T.; IYER, L. **Land rights and economic development: evidence from Viet Nam.** MIT, 2002. Mimeografado.

FEDER, G. *et al.* **Land policies and farm productivity in Thailand.** Baltimore: Johns Hopkins University Press, for the World Bank, 1988.

FERNANDES, E. Principles, bases and challenges of the national programme to support sustainable urban land regularization in Brazil. In: HUCHZERMEYER, M.; KARAM, A. (Ed.). **Informal settlements. A perpetual challenge?** University of Cape Town Press, 2006. p. 62-83.

FIELD, E. **Entitle to work: urban property rights and labor supply in Peru.** Princeton: Princeton University, July 2003a. Mimeografado.

_____. **Fertility reponses to land titling: the roles of ownership security and the distribution of household assets.** Princeton: Princeton University, 2003b. Mimeografado.

_____. Property rights and investment in urban slums. **Journal of European Economic Association**, p. 279-290, Apr./May 2005.

_____.; TORERO, M. **Do property titles increase credit access among the urban poor? Evidence from a nationwide titling program.** Princeton: Princeton University, Mar. 2006. Mimeografado.

FRIEDMAN, J.; JIMENEZ, E.; MAYO, S. The demand for tenure security in developing countries. **The Journal of Development Economics**, v. 29, n. 2, p. 185-198, 1988.

GALEANA, F. **Explaining the demand for land titling in Mexico: credit and urbanization in the ejido sector.** Stanford University Center for Latin American Studies, 2004. Mimeografado.

GALIANI, S.; SCHARGRODSKY, E. **Effects of land titles on child health.** Washington: Inter-American Development Bank, July 2004 (Research Network Working Paper, R-491).

_____. **Effects of land titles.** Buenos Aires: Universidad Torcuato Di Tella, jan. 2007. Mimeografado.

HECKMAN, J.; ROBB, R. Alternative methods for evaluating the impact of interventions: an overview. **Journal of Econometrics**, v. 30, p. 239-267, Apr. 1985.

_____.; ICHIMURA, H.; TODD, P. Matching as an econometric evaluation estimator. **Review of Economic Studies**, v. 65, n. 2, p. 261-294, Apr. 1998.

- JIMENEZ, E. Tenure security and urban squatting. **The Review of Economics and Statistics**, v. 66, n. 4, p. 556-567, Nov. 1984.
- LANJOUW, J. O.; LEVY, P. I. Untitled: a study of formal and informal property rights in urban Ecuador. **The Economic Journal**, v. 112, p. 986-1.019, Oct. 2002.
- LOPEZ, R.; ROMANO, C. **Rural poverty in Honduras: asset distribution and liquidity constraints**. 1997. Mimeografado.
- MICELI, T.; SIRMANS, C. F.; KIEYAH, J. The demand for land title registration: theory with evidence for Kenya. **American Law of Economics Review**, v. 3, n. 2, p. 275-287, 2001.
- MIGOT-ADHOLLA, S. E. *et al.* Indigenous land rights systems in Sub-Saharan Africa: a constraint on productivity? **World Bank Economic Review**, v. 5, n. 1, p. 155-175, 1991.
- PAYNE, G.; DURAND-LASSERVE, A.; RAKODI, C. **Social and economic impacts of land titling programmes in urban and peri-urban areas: a review of the literature**. World Bank Urban Research Symposium. Washington D.C., May 2007.
- PINCKNEY, T. C.; KIMUYU, P. Land tenure reform in East Africa: good, bad or unimportant? **Journal of African Economies**, v. 3, n. 1, p. 1-28, 1994.
- PLACE, F.; OTSUKA, K. Land tenure systems and their impacts on agricultural investments and productivity in Uganda. **Journal of Development Studies**, v. 38, n. 6, p. 105-124, 2002.
- RESENDE, A. C. C.; OLIVEIRA, A. M. H. C. de. Avaliando resultados de um programa de transferência de renda: o impacto do Bolsa-Escola sobre os gastos das famílias brasileiras. **Estudos Econômicos**, São Paulo, v. 38, n. 2, p. 235-265, abr./jun. 2008.
- ROSENBAUM, P. R. **Observational studies**. New York: Springer, 2002.
- RUBIN, D. B., THOMAS, N. Matching using estimated propensity scores: relating theory to practice. **Biometrics**, v. 52, p. 249-264, 1996.
- SILVA, A. M. A. da; RESENDE, G. M.; SILVEIRA NETO, R. da M. Eficácia do gasto público: uma avaliação do FNE, FNO e FCO. **Estudos Econômicos**, São Paulo, v. 39, n. 1, mar. 2009.
- UN-HABITAT. Land tenure, housing rights and gender – national and urban framework. **Law, Land Tenure and Gender Review Series: Latin America, Nairobi**, 2005.
- VIAL, A.; CAVALLIERI, F. **O efeito da presença governamental sobre a expansão horizontal das favelas do Rio de Janeiro: os pousos e o Programa Favela-Bairro**. Rio de Janeiro: Prefeitura da Cidade do Rio de Janeiro, Instituto Pereira Passos, maio 2009 (Coleção Estudos Cariocas, n. 20090501). Disponível em: <armazemdedados.rio.rj.gov.br>
- WOOLDRIDGE, J. M. **Introdução à econometria: uma abordagem moderna**. 2. ed. Thompson Learning, 2000.
- _____. **Econometrics analysis of cross section and panel data**. Cambridge: The MIT Press, 2002.

(Originais submetidos em dezembro de 2010. Última versão recebida em abril de 2011.
Aprovada em abril de 2011.)

TRABALHO MATERNO E DESEMPENHO EDUCACIONAL DAS CRIANÇAS: UMA ANÁLISE DA PROBABILIDADE DE APROVAÇÃO ESCOLAR*

Juliana Maria de Aquino**

Elaine Toldo Pazello***

O objetivo deste estudo é analisar o impacto do trabalho materno sobre a probabilidade de aprovação escolar das crianças brasileiras com idade entre 10 e 14 anos. Para tanto, foram utilizadas duas subamostras da Pesquisa Mensal de Emprego (PME), do Instituto Brasileiro de Geografia e Estatística (IBGE), referentes aos períodos 1986-1995 e 2002-2006. Os resultados obtidos mostram que, para a amostra mais antiga, o fato de a mãe ingressar no mercado de trabalho não teve impacto sobre a probabilidade de aprovação escolar da criança. Já na amostra mais recente, os resultados foram mais expressivos: o impacto do trabalho materno apresentou uma magnitude de $-7,1$ pontos percentuais. Outro aspecto interessante é que o impacto negativo é sempre maior para as mães mais escolarizadas. A divergência de resultados entre os subperíodos pode ser decorrente: *i*) das mudanças metodológicas da PME; *ii*) das diferenças em termos de variação de renda das famílias entre os períodos; *iii*) das mudanças das “regras” relativas à aprovação/reprovação nas escolas; ou, por fim, *iv*) do aumento da escolarização das mães entre os períodos analisados. Estudos futuros podem buscar identificar se de fato estes elementos são importantes e, em caso positivo, qual a contribuição de cada um deles no resultado final.

1 INTRODUÇÃO

Nos últimos anos, o tema educação tem sido alvo de muitos estudos, certamente em função de sua importância sobre os diversos aspectos econômicos e sociais dos indivíduos e dos países. No Brasil, tendo em vista os números nada animadores dos indicadores de desempenho escolar, estudos voltados para o entendimento dos determinantes do desempenho educacional ganham cada vez mais espaço e se fazem necessários para formulações de políticas educacionais. Este é o contexto no qual este trabalho se insere.

Diversas pesquisas apontam o *background* familiar como um dos principais determinantes do desempenho educacional.¹ Em geral, os estudos utilizam para *background* familiar as características dos pais, em termos de escolaridade e idade.

* As autoras agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro, e aos pareceristas anônimos por suas sugestões.

** Doutoranda em Economia pela Escola Superior de Agricultura Luiz de Queiroz (ESALQ)/Universidade de São Paulo (USP).

*** Do Departamento de Economia da Faculdade de Economia e Administração e Contabilidade de Ribeirão Preto (FEA-RP)/USP.

1. Exemplos de estudos com tais resultados para o Brasil são: Barros *et al.* (2001), Albernaz, Ferreira e Franco (2002) e Felício e Fernandes (2005). Sobre o debate – ainda em aberto – a respeito do papel da família *versus* da escola no desempenho escolar ver, dentre outros, Hanushek (2003, 2007).

Este artigo, no entanto, foca especificamente o papel da presença da mãe no aprendizado da criança. A ideia é analisar o impacto do ingresso da mãe no mercado de trabalho sobre o desempenho escolar de seus filhos, especificamente sobre a probabilidade de aprovação das crianças. Para tanto, serão utilizados dados obtidos a partir de duas subamostras da Pesquisa Mensal de Emprego (PME), do Instituto Brasileiro de Geografia e Estatística (IBGE), dos anos de 1986 a 1995 e de 2002 a 2006. O trabalho contribui com a literatura, primeiro, porque analisa um atributo da família pouco explorado nos estudos brasileiros sobre o tema e segundo por utilizar os dados da PME, os quais proporcionarão um trabalho diferenciado, permitindo que se trabalhe com um painel de indivíduos.

A participação no mercado de trabalho reduz, em certo número de horas, o tempo de dedicação com o cuidado aos filhos. E é a partir desse mecanismo que o emprego materno pode ter efeitos sobre os resultados escolares das crianças. Na literatura, o impacto do trabalho materno sobre o desempenho educacional da criança tem apresentado resultados contraditórios. Essa diversidade, provavelmente, decorre da perspectiva ambígua do trabalho materno, isto é, se por um lado a ausência da mãe no domicílio pode reduzir o nível de controle, de orientação e de monitoração dados à criança, por outro, o trabalho materno pode estar associado a uma renda familiar maior, compensando, assim, a redução no tempo gasto com a criança.² A interpretação econômica dessa perspectiva enfatiza o *trade-off* com o qual os pais se deparam, ou seja, eles devem decidir sobre a combinação entre os recursos monetários e temporais que serão utilizados na produção de qualidade da atenção dada à criança.

As pesquisas que abordam tal relação apresentam resultados sensíveis às características da amostra, às medidas escolhidas para representar o emprego materno e o desempenho escolar da criança, e, principalmente, à escolha das variáveis de controle. Estudos americanos como os de Milne *et al.* (1986) e Krein e Beller (1988) encontraram efeitos restritivos do trabalho materno sobre o desempenho educacional das crianças. Todavia, os resultados se limitam a crianças de cor branca. Por outro lado, o estudo de Haveman, Wolfe e Spaulding (1991) aponta o trabalho materno como um determinante positivo para a conclusão do ensino médio. Contudo, o impacto depende do momento em que ele ocorre: o fato de a mãe trabalhar e, desta forma, contribuir com a renda familiar, tem impacto positivo somente quando o evento ocorre na adolescência dos alunos. Em contrapartida, o efeito positivo é reduzido quando o trabalho materno ocorre durante a fase de infância do filho. Resultados semelhantes ao anterior também foram encontrados por Blau e Grossberg (1992).

Ainda se tratando da literatura internacional, estudos mais recentes têm enfatizado os efeitos adversos com o retorno precoce da mãe ao mercado de trabalho

2. Ver, a este respeito, Haveman e Wolfe (1995).

após o nascimento de um filho. Waldfogel *et al.* (2002) encontraram efeitos restritivos do trabalho materno durante o primeiro ano de vida da criança, os quais se mostraram negativamente mais intensos entre as mães que trabalhavam em tempo integral.³ O estudo de Gregg *et al.* (2005), a partir de dados do Reino Unido, também mostra que o trabalho materno em jornada integral tem efeitos adversos, porém de pequena magnitude. No entanto, os autores observam que os efeitos tendem a ser piores conforme aumenta o nível de escolaridade das mães. Por outro lado, são menos perversos para as crianças provindas de famílias monoparentais.

No Brasil, a relação entre trabalho materno fora de casa e desempenho educacional dos filhos tem sido pouco explorada. Na verdade, o único trabalho de que se teve conhecimento foi o realizado por Silva e Hasenbalg (2000), que avalia a evolução das desigualdades educacionais e as mudanças nos determinantes extraescolares do desempenho escolar. A partir dos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD), do IBGE, referentes aos anos de 1976, 1986 e 1998, os autores analisam um modelo de realização escolar no qual o nível de escolarização atingido pelos jovens de 15 a 18 anos é predito por fatores indicativos da situação socioeconômica, do capital cultural e do capital social acessível a suas famílias. Nesse modelo, o trabalho materno representa uma das *proxies* de capital social, dado que a ausência da mãe no domicílio reduz a quantidade de tempo disponível para o relacionamento com os filhos, diminuindo, assim, o montante de capital social dentro da família. Os resultados evidenciaram um comportamento não esperado para a variável de trabalho materno: o efeito desta variável se transformou de negativo, em 1976, para positivo em 1998.

Em suma, a resenha da literatura mostrou que o efeito do trabalho materno sobre os resultados escolares da criança varia conforme o corte amostral, ou seja, não existe consenso sobre esse resultado na literatura. Como pôde ser observado, raça, tipo de composição familiar, período de vida da criança no qual o evento ocorre e variáveis maternas como jornada de trabalho e escolaridade são fatores que influenciam fortemente o sinal e até mesmo a magnitude do efeito observado. Além disso, também se pôde constatar escassez de estudos brasileiros que tratam dessa relação, o que dá à pesquisa proposta a oportunidade de trazer resultados para um melhor entendimento da relação entre trabalho materno e desempenho educacional no Brasil.

Este estudo está organizado em mais quatro seções, além desta introdução. A segunda e a terceira seções descrevem, respectivamente, os dados e a metodologia. A quarta seção apresenta os resultados obtidos. A última seção é dedicada às considerações finais.

3. Resultados semelhantes foram obtidos por Ruhm (2002).

2 BASE DE DADOS E SELEÇÃO DA AMOSTRA

A base de dados utilizada nesse estudo foi construída a partir da PME. Essa pesquisa, iniciada em 1980, tem como principal objetivo produzir indicadores do mercado de trabalho. Sua periodicidade é mensal e abrange as regiões metropolitanas (RMs) de Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo e Porto Alegre. O diferencial dessa pesquisa é que ela trabalha com painéis amostrais rotativos de domicílios. O domicílio compreendido na amostra é entrevistado por um período de quatro meses consecutivos, é retirado da amostra por oito meses e retorna a ela exatamente após um ano a contar da primeira entrevista, sendo entrevistado por mais quatro meses.

Em 2001, a PME passou por um processo de revisão metodológica de forma a atualizar sua cobertura temática e se adequar às recomendações internacionais. As mudanças foram de porte significativo, de forma que há uma “quebra” na série histórica da pesquisa: de 1980 a 2001 tem-se dados de acordo com a metodologia antiga da PME, e de 2002 em diante os dados seguem a metodologia nova da PME. Optou-se, nesse estudo, por utilizar os dados tanto da PME “antiga” quanto da PME “nova”. Tal decisão se apoia nas vantagens exclusivas do uso de cada uma delas. A PME “antiga” abrange um maior período de anos, possibilitando, assim, a formação de amostras mais representativas. Além disso, o período compreendido entre 1986 e 1995 pode ajudar a suprimir eventuais problemas que as políticas educacionais de progressão continuada possam ter em relação à probabilidade de aprovação das crianças, dado que o desdobramento do ensino fundamental em ciclos foi facultado pela Lei de Diretrizes e Bases da Educação Nacional (LDB) de 1996. É por essa razão, também, que se optou por trabalhar apenas com anos de 1986 a 1995 da PME antiga. Por outro lado, a PME nova torna possível a incorporação de uma importante variável referente ao passado ocupacional da mãe, que não estava presente no questionário antigo da PME. É importante enfatizar que a inexistência de informações na PME antiga acerca do grau de *attachment* das mães ao mercado de trabalho resulta em uma considerável limitação dessa base de dados na seleção da amostra de interesse. É possível, por exemplo, que a sazonalidade⁴ seja importante e, por não se ter informações sobre o passado ocupacional da mãe, limita-se o controle dessa variável. Na amostra da PME nova esse problema se reduz, já que se consegue obter informações sobre a presença materna no mercado de trabalho pelo menos nos últimos 12 meses.

Embora a PME possibilite uma gama muito rica para estudos longitudinais, ainda existe uma subutilização desta pesquisa, segundo Ribas e Soares (2008),

4. Para se ter uma ideia da sazonalidade no mercado de trabalho, verificou-se, para alguns anos da PME, a proporção mensal de mulheres, com mais de 18 anos, que estavam fora do mercado de trabalho. Os resultados mostram pouca variação nessa proporção ao longo dos meses.

principalmente, em função das dificuldades técnicas para a (re)construção do painel de indivíduos. De qualquer forma, há diversos estudos que usam a PME como base de dados. Os trabalhos em geral são voltados para o tema de mercado de trabalho propriamente: Menezes-Filho e Picchetti (2000) analisam os determinantes da duração do desemprego; Fernandes e Chahad (2002) analisam o impacto do seguro-desemprego no Brasil sobre as probabilidades de transição no mercado de trabalho; Gonzaga, Menezes-Filho e Camargo (2003) analisam o impacto da redução da jornada de trabalho (de 44 horas para 40 horas, da Constituição de 1988) sobre engajamento e salário; Fernandes e Felício (2005) analisam o efeito “trabalhador adicional”, isto é, se há uma resposta da esposa em termos de participação no mercado de trabalho quando o marido perde o emprego; Curi e Menezes-Filho (2006) utilizam o painel da PME para discutir as diferenças entre os mercados de trabalho formal e informal. Mas há também exemplos de estudos que exploram outras temáticas: Leon e Menezes-Filho (2002) analisam os determinantes das taxas de repetência, avanço e evasão escolar usando o painel da PME; Cavalieri (2000) investiga o impacto do trabalho infantil sobre resultados escolares; e Machado, Ribas e Penido (2007) analisam a relação entre mobilidade social e as transições no mercado de trabalho. O presente estudo se encaixa neste segundo grupo de trabalhos.

A seleção da amostra de interesse seguiu os procedimentos adotados em Cavaliere (2000).⁵ Primeiramente, só foram incluídas as famílias cuja primeira entrevista ocorreu entre os meses de março e julho do ano em questão. Essa escolha decorre do fato de os dois primeiros meses do ano corresponderem às férias escolares e, assim, não seria possível identificar se a criança estaria frequentando a escola ou não. Por outro lado, nos meses de agosto a dezembro, parte do ano letivo já se foi e, desta forma, o desempenho escolar da criança já pode estar comprometido por outras razões que podem não estar relacionadas ao trabalho materno. O trabalho utiliza apenas as cinco primeiras entrevistas.⁶ Assim, por exemplo, se a primeira entrevista ocorre em março, acompanha-se a família, em abril, maio e junho desse mesmo ano e em março do ano seguinte.

Dessas famílias, só foram incluídas na amostra crianças com idade entre 10 e 14 anos, que estivessem frequentando a escola quando foram entrevistadas pela primeira vez e que permaneceram estudando durante todas as cinco entrevistas. Além das restrições impostas anteriormente, somente fizeram parte da amostra

5. Neste trabalho, o objetivo da autora foi o de analisar a relação entre trabalho infantil e desempenho escolar. Para isso também utilizou dados da PME. O desenho usado pela autora para identificar o efeito causal do trabalho infantil foi adaptado para o presente estudo.

6. No painel da PME, os entrevistadores voltam no mesmo domicílio e, assim, não necessariamente na mesma família visto que ela pode mudar de domicílio entre as entrevistas. Dessa forma, é necessário utilizar um identificador para se assegurar de que a família observada seja a mesma em todas as entrevistas. Neste estudo, a identificação das famílias foi feita a partir da data de nascimento da mãe. Por esse critério, famílias cuja mãe não informou ou informou erroneamente a data de nascimento tiveram de ser excluídas da amostra.

aquelas crianças cujas mães estiveram fora do mercado de trabalho quando da primeira entrevista, ou seja, crianças cujas mães eram inativas.

Vale ressaltar que foram excluídas da amostra famílias cujas mães não responderam às questões referentes à data de nascimento, visto que inviabilizariam a tentativa de garantir que as mães do domicílio entrevistado fossem as mesmas em todas as cinco entrevistas. No final da seleção, o mesmo procedimento foi adotado para as crianças da amostra. Foram excluídas também as crianças que não pertenciam ao grupo de tratamento e nem ao de controle. A composição de tais grupos será explicitada a seguir. Finalmente, foram retiradas aquelas famílias que tiveram seu *status* conjugal alterado, assim como as crianças cuja diferença de anos de escolaridade entre os dois anos analisados fosse diferente de zero ou 1. A tabela 1 apresenta uma síntese cronológica das etapas realizadas na composição da amostra de interesse, assim como o número de observações presentes em cada fase da seleção.

TABELA 1

Etapas da seleção amostral

Número da etapa e sua designação	Número de crianças com idade entre 10 e 14 anos	Número de crianças com idade entre 10 e 14 anos
	Amostra de 1986-1995	Amostra de 2002-2006
1ª etapa: foram selecionadas somente as crianças com idade entre 10 e 14 anos cuja primeira entrevista ocorreu entre os meses de março a julho.	74.379	17.648
2ª etapa: foram excluídas as crianças cuja mãe não respondeu as questões referentes à data de nascimento.	73.252	15.580
3ª etapa: permaneceram na amostra somente as crianças que estavam frequentando a escola quando da primeira entrevista.	69.016	15.235
4ª etapa: foram selecionadas somente as crianças cuja mãe estivesse fora do mercado de trabalho quando da primeira entrevista.	36.455	6.371
5ª etapa: foram excluídas as crianças cuja mãe não participou de todas as cinco entrevistas.	27.631	4.481
6ª etapa: permaneceram na amostra somente as crianças cuja mãe fosse a mesma em todas as cinco entrevistas (verificação a partir da data de nascimento).	27.446	3.444
7ª etapa: foram excluídas aquelas crianças que não pertenciam ao grupo de tratamento nem ao de controle.	20.308	2.617
8ª etapa: foram selecionadas somente as crianças cuja mãe não teve o <i>status</i> conjugal alterado.	20.170	2.582
9ª etapa: permaneceram na amostra apenas as crianças que estavam frequentando a escola em todas as cinco entrevistas.	18.324	2.143
10ª etapa: foram excluídas as crianças cuja diferença de anos de escolaridade entre os dois anos analisados fosse diferente de zero ou 1.	17.235	1.829
11ª etapa: foram selecionadas somente as crianças que responderam as questões referentes à data de nascimento.	16.998	1.775
12ª etapa: permaneceram na amostra as crianças que fossem as mesmas em todas as cinco entrevistas (verificação a partir da data de nascimento).	15.371	1.464

Fonte: PME. Elaboração própria.

Da tabela 1 é interessante destacarmos a questão do atrito que permeia o painel da PME. Como já colocado, a PME volta ao domicílio, mas nem sempre a família que morava continua sendo a mesma. Ribas e Soares (2008) fazem uma discussão bastante rica sobre a taxa de atrito nas PMEs antiga e nova. Segundo os cálculos destes autores, a taxa de atrito entre a primeira e a quinta entrevista chega a 26% e 55% na PME antiga e na nova, respectivamente. Ou seja, 26% e 55% dos indivíduos são “perdidos” entre a primeira e a quinta entrevista, tanto na PME antiga quanto na PME nova. A comparação entre as etapas 4 e 6 da tabela 1 nos permite calcular números comparáveis aos de Ribas e Soares (2008). Para a PME antiga, a taxa de atrito deste trabalho é de 24,7%; para a PME nova, a taxa de atrito é de 46%. Ou seja, os números são muito próximos aos obtidos por estes autores.⁷

A partir dessa amostra, foram identificados dois grupos. O primeiro grupo, denominado grupo de controle (a seção metodológica explicará essa nomenclatura), é composto por crianças cujas mães não participaram do mercado de trabalho em nenhuma das cinco entrevistas observadas, isto é, eram inativas na primeira entrevista e continuaram inativas durante as outras quatro entrevistas. O outro grupo, que é o grupo de tratamento, é formado por crianças cujas mães passaram a fazer parte da população economicamente ativa (PEA) em um dos três meses subsequentes à primeira entrevista, e continuaram participando nos meses em que foram entrevistadas nesse mesmo ano e no ano seguinte, ou seja, até a quinta entrevista. Um exemplo pode ajudar a entender. Suponha que a mãe tenha sido observada pela primeira vez em março e que ela tenha entrado no mercado em maio (isto é, na terceira entrevista). Se este foi o caso, para fazer parte desse grupo, ela deve ter permanecido ativa em junho daquele ano e também em março do ano subsequente (ou seja, na quinta entrevista). Desta maneira, a amostra do período 1986-1995 foi composta por 13.748 observações, das quais 1.627 representam o grupo de tratamento. E para o segundo período, 2002-2006, a amostra formada compreende 1.398 observações, sendo o grupo de tratamento composto por 211 crianças.⁸

3 METODOLOGIA

A avaliação do impacto da entrada da mãe no mercado de trabalho com relação à probabilidade de aprovação escolar da criança se resume na análise de como teria sido o desempenho escolar das crianças cujas mães começaram a participar do mercado de trabalho, caso elas não tivessem começado a participar. No entanto,

7. Os autores propõem um algoritmo para melhorar o emparelhamento de indivíduos nos painéis da PME – usar outras variáveis que não a data do nascimento. Há um ganho razoável para a PME nova, mas as taxas de atrito ainda continuam sendo bastante grandes.

8. Vale ressaltar que o número de observações da amostra final difere do número de observações da última etapa da seleção amostral, apresentada na tabela 1, devido à presença de *missing* nas variáveis explicativas utilizadas nos modelos.

dada a impossibilidade de se observar aquelas mães que participaram do mercado, porém na situação de não terem participado, a solução foi criar um grupo de controle a fim de representar tal situação, de maneira que a avaliação do contrafactual exposto seja alcançada. Na linguagem de avaliação, o “tratamento”⁹ é a entrada da mãe no mercado de trabalho; conseqüentemente, o “grupo de tratamento” é formado pelas crianças cujas mães começaram a trabalhar e o “grupo de controle” é formado pelas crianças cujas mães permaneceram fora do mercado de trabalho.

Serão apresentados, no decorrer do texto, impactos estimados tanto por Mínimos Quadrados Ordinários (MQO) como por *P propensity Score Matching* (PSM). As estimativas obtidas por MQO não ponderam as observações de acordo com seu grau de semelhança com o grupo de tratados; o pareamento, por outro lado, leva em conta o grau de semelhança, atribuindo maior peso às unidades do grupo de controle que são mais próximas às unidades tratadas. Os dois métodos, no entanto, se baseiam na mesma hipótese de identificação, especificamente, de “seleção por observáveis” ou *unconfoundedness* (ignorabilidade do tratamento). Ou seja, por hipótese, condicional num vetor de características X , não há nada sistemático que leve uma observação a ser tratada ou não. Dito de outra forma, os determinantes não observáveis do desempenho escolar da criança, controladas as características X , não são correlacionados com o ingresso da mãe no mercado de trabalho (isto é, com o tratamento). Segue a derivação formal do exercício de interesse deste estudo.

Seja Y_{1i} o resultado, em termos de aprovação escolar, da criança i cuja mãe começou a trabalhar e Y_{0i} o resultado da criança i cuja mãe continuou fora do mercado de trabalho. Assim, o efeito da entrada da mãe no mercado de trabalho é $\pi_i = Y_{1i} - Y_{0i}$ e, o impacto médio da entrada da mãe no mercado de trabalho sobre tais crianças é $\Pi = E[Y_{1i} - Y_{0i} | T_i = 1]$, onde $T_i = 1$ indica que a mãe da criança i entrou no mercado de trabalho e $T_i = 0$ indica que a mãe não entrou no mercado de trabalho. Entretanto, como já colocado, não é possível verificar a mesma criança nessas duas situações. Nesse caso, a ideia é eleger um grupo de comparação que represente o grupo tratado na ausência do tratamento.

O objetivo é, então, estimar a diferença entre os resultados escolares, em termos de aprovação, das crianças cujas mães passaram a participar do mercado de trabalho e os resultados das crianças cujas mães não entraram no mercado de trabalho, de forma que:

$$E[Y_{1i} | T_i = 1] - E[Y_{0i} | T_i = 0] = \Pi + \{E[Y_{0i} | T_i = 1] - E[Y_{0i} | T_i = 0]\} \quad (1)$$

9. Outras configurações de tratamento serão testadas no decorrer do texto, na análise de robustez.

O termo (1) é uma medida do viés de seleção, representando o fato de os resultados escolares das crianças cujas mães não entraram no mercado não corresponderem exatamente aos resultados, em termos de aprovação escolar, das crianças cujas mães passaram a trabalhar, caso tais mães não tivessem entrado no mercado de trabalho. Em outras palavras, para identificar Π por meio da comparação com um grupo de crianças não tratadas, seria necessário garantir que $E[Y_{0i}|T_i = 1] - E[Y_{0i}|T_i = 0] = 0$. Esta última condição seria atendida, caso a designação das crianças entre tratamento e controle ($T = 0,1$) tivesse ocorrido de forma aleatória. No presente exercício, entretanto, certamente o processo de seleção não foi aleatório.

O que significa dizer que a designação entre tratamento e controle não é aleatória? Por exemplo, é provável que exista uma sobrerrepresentação de mães mais educadas entre aquelas que decidem entrar no mercado de trabalho. Assim, as crianças no grupo de tratamento devem ter, em média, mães mais educadas. A simples comparação das médias de aprovação entre os grupos pode evidenciar uma probabilidade maior de aprovação entre as crianças cujas mães começaram a trabalhar. No entanto, esta diferença, provavelmente, se deve, em parte, ao fato de as mães do grupo de tratamento apresentarem maior nível educacional, visto que um dos efeitos mais bem evidenciados na literatura de educação é o efeito positivo da escolaridade da mãe sobre o desempenho escolar da criança.

Como discutido no início desta seção, a hipótese de identificação adotada é a de que mães que possuam as mesmas características observáveis teriam a mesma probabilidade de ingresso no mercado de trabalho. Isto significa assumir que $(Y_{0i}, Y_{1i} \perp T_i) | X_i$ e $E[Y_{0i} | X_i, T_i = 1] = E[Y_{0i} | X_i, T_i = 0]$, onde \perp denota independência e X o vetor de características observáveis.¹⁰ Portanto, a plausibilidade desta hipótese depende da escolha de X . Evidentemente, isso não garante que não haja diferenças não observáveis entre os grupos; a ideia é incorporar um conjunto de observáveis que minimizem essa possibilidade. De qualquer forma, assumindo-se esta hipótese, obtém-se:

$$\Pi_X = E[Y_i | X_i, T_i = 1] - E[Y_i | X_i, T_i = 0] \quad (2)$$

onde Π_X é o impacto médio da entrada da mãe no mercado de trabalho sobre a probabilidade de aprovação da criança, condicional em X .

10. A hipótese de identificação no caso das estimativas por PSM: se garantirmos que o viés de seleção é eliminado por se condicionar às estimativas no vetor de características X , segundo Rosenbaum e Rubin (1983), vale também a independência condicional, dado o propensity score $P(X)$.

A variável dependente empregada neste estudo é a probabilidade de aprovação da criança. Essa variável foi obtida por meio da análise dos anos de estudo completos entre os dois anos consecutivos em que cada criança é observada na amostra. Desta forma, foram consideradas aprovadas as crianças cuja diferença de anos de estudo entre a quinta e a primeira entrevista fosse igual a 1. Naturalmente, foram identificadas como repetentes as crianças para as quais tal diferença foi igual a zero. Em função da característica binária dessa variável, os resultados econométricos são obtidos a partir de modelos *logit*.

Com relação ao vetor de características X , foram incluídas as seguintes variáveis:

- relativas às crianças: gênero, idade, atraso escolar¹¹ e participação no mercado de trabalho (trabalho infantil);
- relativas à família da criança: se a criança residia com ambos os pais, número de irmãos com idade igual ou inferior a 14 anos, número de irmãos maiores de 15 anos, renda familiar *per capita* na primeira entrevista¹² e variação anual da renda familiar *per capita* entre a quinta e a primeira entrevista (sem incluir a renda oriunda do trabalho da mãe caso esta tenha entrado no mercado de trabalho);
- relativas ao pai da criança: escolaridade e uma *dummy* para identificar se o pai era empregado em tempo integral;
- relativas à mãe da criança: idade, escolaridade e uma *dummy* para identificar se a mãe tinha trabalhado nos últimos 12 meses anteriores à primeira entrevista; e
- foram incluídas também *dummies* para cada uma das RMs e *dummies* para os diferentes anos e meses da PME.

Dentre as variáveis relativas à criança, destaque deve ser dado ao atraso escolar e à condição de trabalho. A variável atraso escolar deve estar captando o efeito fixo individual de cada criança que não é observável nos dados, visto que é basicamente uma variável do “histórico escolar” de cada criança. Isto é, representa as atitudes passadas de cada criança com relação à escola: por exemplo, o atraso pode refletir dificuldade com os estudos ou até mesmo falta de interesse. As evidências empíricas mostram uma relação negativa entre atraso e desempenho escolar, logo, se houver

11. Para o cálculo da variável de atraso escolar, foi necessário construir primeiro a escolaridade ideal para as crianças da amostra. De acordo com os critérios utilizados, estaria em situação “adequada” em termos de escolaridade, por exemplo, o estudante que completasse 10 anos até julho de um determinado ano e que estivesse, naquele ano, frequentando a 4ª série do ensino fundamental, ou seja, que tivesse três anos completos de estudo. Raciocínio análogo pode ser aplicado para as demais idades. Calculada a escolaridade ideal, a variável atraso escolar se resume na diferença entre os anos de estudo completos que seria esperado para sua idade e os anos de estudo efetivamente concluídos pela criança. Medidas semelhantes dessa variável foram utilizadas em Horowitz e Souza (2004) e Machado e Gonzaga (2007).

12. Na PME, a renda familiar inclui apenas a renda oriunda do mercado de trabalho (salários, especificamente).

qualquer correlação entre essa variável e a inserção da mãe no mercado de trabalho, o impacto estimado seria enviesado sem o seu controle.

A variável trabalho infantil deve estar correlacionada com as condições financeiras da família. Quanto menor a renda familiar maior a probabilidade de a criança estar trabalhando, o que pode afetar seu desempenho escolar, já que o trabalho reduz o tempo disponível para o estudo extraescolar e ainda pode prejudicar o aluno em sala de aula, em virtude do cansaço causado pelo trabalho.¹³ Por outro lado, uma menor renda familiar pode levar também a uma participação maior de outros membros da família, inclusive da mãe. Assim, é importante condicionar as estimativas a esta variável, porque caso contrário seu efeito negativo seria captado pela variável de interesse, isto é, pela variável que identifica o tratamento.

Com relação às características da família, o destaque fica para a “variação anual da renda familiar *per capita*”. Suponha que haja entre a primeira e a quinta entrevista um choque negativo sobre a renda da família, causado, por exemplo, pela perda do emprego do pai. Isto pode levar parte das mães para o mercado de trabalho. Por outro lado, esta queda da renda familiar pode também afetar o desempenho educacional das crianças. Assim, se não houver um controle para esse choque, parte de seu efeito será erroneamente identificado como um efeito negativo da entrada da mãe no mercado de trabalho sobre o desempenho escolar das crianças. É por esta razão que a amostra foi restrita apenas às famílias que não tiveram alteração de *status* conjugal entre as entrevistas.

Quanto às características paternas, a ideia é que essas variáveis – por exemplo, a jornada de trabalho dos pais – sejam relacionadas tanto com a probabilidade de aprovação da criança como com a entrada da mãe no mercado de trabalho.

Por fim, com relação às características das mães, o destaque será para a variável de passado ocupacional. Pode ser que o fato de a mãe ter começado a participar do mercado de trabalho, e dessa forma ter de se ausentar do ambiente domiciliar, tenha efeitos restritivos sobre o desempenho escolar das crianças somente num primeiro momento, até que as crianças se adaptem a essa nova rotina familiar. Dessa forma, quanto mais recorrente para a mãe for o evento “entrar no mercado de trabalho”, menos custoso será para a criança, já que esta estará mais habituada aos momentos de ausência da mãe. Esta informação sobre o passado ocupacional da mãe, no entanto, não existe no questionário da PME antiga e, assim, somente estará presente nos exercícios referentes à amostra do período 2002-2006. Já as variáveis de idade e escolaridade são características que refletem a produtividade das mulheres e que, portanto, devem estar correlacionadas com a inserção da mãe no mercado de trabalho.

13. A relação negativa entre renda familiar e probabilidade de a criança trabalhar é bem conhecida na literatura, podendo ser observado, por exemplo, no estudo de Cavaliere (2000) e Kassouf (1999).

Com relação aos demais controles, para os quais não foram feitos comentários específicos, a ideia é sempre a mesma: são variáveis que afetam o desempenho escolar das crianças e que estão correlacionadas com a entrada da mãe no mercado de trabalho, ou seja, com a designação das crianças entre tratamento e controle. Assim, se tais controles não forem considerados haverá correlação entre a variável de interesse e o erro, enviesando as estimativas.

Evidentemente, não se pode garantir que não haja diferenças não observáveis entre os grupos – não há como testar tal hipótese. Na tentativa de minimizar essa possibilidade, buscou-se construir a amostra de forma a tornar o exercício mais próximo possível de um experimento. Exigiu-se, por exemplo, que no início todas as mães estivessem fora do mercado de trabalho e que não tivessem alteração de seu *status* conjugal durante todo o tempo em que foram observadas. Além dessas tentativas de homogeneizar os grupos, serão incorporadas importantes variáveis entre os controles, como, por exemplo, a variável que capta o atraso escolar da criança no primeiro momento, ou seja, quando todas as crianças ainda contam com a presença integral de suas mães em casa. Com relação às diferenças não observáveis entre os grupos, pode ser que fatores relacionados a crises familiares, como brigas e discussões entre os membros da família, estejam ao mesmo tempo levando a mãe ao mercado de trabalho, na fuga de um ambiente domiciliar em crise, e ainda estejam determinando o rendimento escolar da criança. Nesse caso, haverá correlação entre a variável de interesse e o erro. Acredita-se, entretanto, que ao controlar a renda familiar esteja-se também captando possíveis efeitos de ambientes familiares em crise. O ideal seria que tivéssemos um instrumento, isto é, uma variável que afetasse os indicadores de aprovação apenas via tratamento. Infelizmente, este não é o caso. Não há como prever a direção do viés que os não observáveis causam sobre o nosso parâmetro de interesse, de forma que não podemos dizer se temos *lower* ou *upper bounds*. Assim, o mais prudente é que se leiam os resultados aqui estimados com a devida cautela.

4 RESULTADOS

Esta seção foi dividida em duas subseções a fim de facilitar a exposição. A primeira apresenta os resultados obtidos a partir dos dados da PME antiga e, a segunda, os obtidos a partir dos dados da PME nova. Para cada um dos exercícios, apresentam-se primeiramente uma descrição dos dados e, em seguida, os resultados propriamente ditos.

4.1 Características descritivas e análise dos modelos *logit* para a amostra de 1986-1995

As características da amostra analisada estão apresentadas na tabela 2. As médias das variáveis são acompanhadas dos respectivos erros-padrão, medida que indica a

precisão das estimativas. Além disso, também foram realizados testes de diferença de médias no intuito de verificar se as diferenças existentes entre os grupos são estatisticamente significativas, ou não.

TABELA 2

Características descritivas dos grupos: média e desvio-padrão – amostra de 1986 a 1995

Variáveis	Grupo de controle		Grupo de tratamento		Diferenças significativas para nível de confiança de 95%
	Média	Desvio-padrão	Média	Desvio-padrão	
Idade da criança	11,91	1,40	11,92	1,39	
Sexo masculino	53%	0,50	50%	0,50	*
Trabalho infantil	2%	0,16	4%	0,20	*
Atraso escolar	1,34	1,38	1,51	1,41	*
Número de irmãos de 0 a 14 anos	1,72	1,49	1,87	1,51	*
Número de irmãos de 15 anos ou mais	0,87	1,30	0,63	1,05	*
Família completa	94%	0,24	89%	0,31	*
Renda familiar inicial <i>per capita</i>	100,10	152,52	64,87	92,92	*
Δ da renda familiar <i>per capita</i>	-22,49	164,53	-16,86	81,48	
% de pais trabalhando em tempo integral	90%	0,30	85%	0,36	*
Educação do pai	4,88	4,54	4,28	3,97	*
Educação da mãe	4,05	3,88	3,81	3,62	*
Idade da mãe	39,71	7,74	37,93	6,64	*
Residência em Salvador	12%	0,32	15%	0,36	*
Residência em Belo Horizonte	22%	0,42	24%	0,43	
Residência em São Paulo	23%	0,42	19%	0,39	*
Residência em Porto Alegre	12%	0,32	12%	0,32	
Residência em Recife	14%	0,34	14%	0,35	
Residência no Rio de Janeiro	17%	0,38	14%	0,35	*
Número de observações	12.121		1.627		

Fonte: PMEs de 1986 a 1995. Elaboração própria.

Nota: *indica diferença significativa.

Em relação às características individuais da criança, pode-se observar que a variável “idade” não apresentou divergência entre o grupo de controle e o grupo de tratamento. As crianças da amostra têm, em média, 12 anos de idade. Por outro

lado, as variáveis referentes ao atraso escolar, gênero e proporção de crianças que participam do mercado de trabalho divergiram entre os dois grupos analisados: as crianças que compõem o grupo de tratamento estão mais distantes da escolaridade ideal e a proporção de trabalhadores infantis neste grupo representa o dobro daquela verificada para o grupo de controle, embora a proporção de meninas seja maior no tratamento relativamente ao controle.

Entre as características familiares – número de irmãos de zero a 14 anos, número de irmãos de 15 anos ou mais, estrutura familiar e renda familiar *per capita* – todas, sem exceção, foram, em média, estatisticamente diferentes entre os grupos. Em média, o grupo de tratamento tem uma quantidade maior de irmãos com idade entre zero e 14 anos e menor quantidade de irmãos maiores de 15 anos. Além disso, tanto a renda como a proporção de famílias completas (pai e mãe presentes) apresentaram valores inferiores para este grupo. Esse último resultado é bastante intuitivo, já que é de se esperar que o chefe de famílias monoparentais, nesse caso as mães, tenha maior necessidade de sair para o mercado em busca de renda. Quanto à variação proporcional da renda, observa-se que enquanto as famílias do grupo de controle tiveram uma variação anual negativa de 22,5%, para o grupo de tratamento a queda representou 26% do valor da renda.¹⁴

Em relação às características dos pais das crianças, a proporção de pais trabalhando em tempo integral é superior no grupo de controle, assim como maior é o nível de escolaridade desses pais para esse grupo.

Quanto às características maternas, é possível observar que tanto a educação como a idade da mãe, fatores produtivos importantes, apresentaram divergências em relação aos grupos em questão: as mães que começaram a participar do mercado de trabalho são mais novas e têm menor escolaridade.

Por último, com relação à região de residência das crianças, observa-se que as diferenças não são significativas para as RMs de Belo Horizonte, Porto Alegre e Recife. Por outro lado, a proporção de crianças do grupo de tratamento é maior na RM de Salvador e menor na de São Paulo.

Em suma, as características da amostra, em sua maioria, apresentaram divergências entre os grupos analisados, destacando-se as variáveis de renda familiar e de trabalho infantil, o que enfatiza a necessidade de condicionar as estimativas nestas características.

A estimação da probabilidade de aprovação da criança em decorrência do ingresso materno no mercado de trabalho foi realizada, inicialmente, a partir de

14. Ainda que o comentário faça referência à variação proporcional da renda das famílias, vale destacar que a tabela descritiva apresenta os valores médios das variações brutas. O cálculo da variação proporcional média da renda, assim como o desvio dessa estimativa, não foi realizado de forma direta devido à existência de valores nulos da renda inicial das famílias.

quatro modelos *logit*. O primeiro modelo inclui somente a *dummy* do grupo de tratamento. No modelo 2 foram inseridas as características das crianças e de sua família. O terceiro modelo compreende, além das variáveis já incluídas nos modelos anteriores, a variável que identifica a renda familiar *per capita* inicial e ainda a variável que controla a variação anual da renda familiar, exclusive a renda da mãe, de forma a controlar eventuais choques que possam ter efeitos sobre o processo educacional da criança. No modelo 4, características paternas foram acrescentadas ao modelo anterior. Para este último modelo, a amostra restringiu-se às crianças provenientes de famílias completas. A tabela 3 apresenta os efeitos marginais, referentes aos modelos mencionados anteriormente, calculados para indivíduos de características médias da amostra.

Em relação à *dummy* de tratamento, a variável de interesse deste estudo, pode-se observar que seus efeitos marginais foram estatisticamente significativos nos três primeiros modelos, mas não foi no último modelo, que é o mais completo. Assim, pode-se dizer que a entrada das mães no mercado de trabalho não tem efeito sobre o desempenho escolar das crianças. Com relação ao último modelo, o qual inclui características paternas da criança, destaca-se o efeito restritivo da *dummy* que identifica pais com emprego em tempo integral. Para essas crianças, a probabilidade de aprovação fica reduzida em 4,1 pontos percentuais (p.p.).

Dentre as variáveis de controle, algumas estimativas merecem ser comentadas. No que tange à variável de defasagem idade-série, nota-se que o aumento do número de anos de atraso escolar da criança tem impacto negativo sobre a chance de aprovação. Em média, o aumento de um ano na defasagem reduz a probabilidade de aprovação em 2,5 p.p.¹⁵ Ainda com relação à defasagem idade-série, é interessante destacar que essa variável pode estar captando o efeito, por exemplo, da variável de escolaridade da mãe. Machado (2005) traz evidências positivas dessa correlação, mostrando que a educação da mãe tem efeitos negativos sobre o atraso escolar de seus filhos. Ainda em relação às características da criança, a probabilidade de aprovação diminui com a idade e é menor entre os meninos.¹⁶ Por outro lado, viver em famílias completas aumenta a probabilidade de a criança ser aprovada, em média, em 3 p.p.

15. Leon e Menezes-Filho (2002) realizaram, também a partir da PME, uma análise dos indicadores e determinantes da reprovação, avanço e evasão escolar condicional à reprovação para 4^{as} e 8^{as} séries do ensino fundamental e 3^{as} séries do ensino médio brasileiro. Esses autores também trouxeram evidências de que menores são as chances de aprovação entre os estudantes com maior atraso escolar.

16. Para a variável discreta "gênero da criança", optou-se por apresentar também as probabilidades preditas nos pontos. Considerando o modelo 4, a probabilidade predita de aprovação é de 0,74 para as meninas e 0,67 para os meninos.

TABELA 3

**Impacto do trabalho materno sobre a probabilidade de aprovação das crianças:
efeitos marginais dos modelos *logit* – amostra de 1986 a 1995**

Variável dependente = "probabilidade de aprovação"	Modelo 1	Modelo 2	Modelo 3	Modelo 4
Crianças cujas mães participaram do mercado	-0,039*** (0,012)	-0,026** (0,013)	-0,023* (0,013)	-0,016 (0,013)
Características da criança				
Trabalho infantil		-0,038 (0,025)	-0,038 (0,025)	-0,042 (0,027)
Anos de atraso escolar		-0,026*** (0,003)	-0,025*** (0,003)	-0,023*** (0,003)
Sexo masculino		-0,064*** (0,008)	-0,064*** (0,008)	-0,067*** (0,008)
Idade		-0,010*** (0,003)	-0,010*** (0,003)	-0,014*** (0,003)
Características da família da criança				
Família completa		0,039** (0,017)	0,032* (0,017)	
Número de irmãos com idade <= 14		-0,009*** (0,003)	-0,007** (0,003)	-0,007** (0,003)
Número de irmãos com idade > 15		-0,003 (0,004)	-0,003 (0,004)	-0,002 (0,004)
Variação anual da renda familiar <i>per capita</i>			0,0001** (0,0001)	0,0001* (0,0001)
Renda familiar <i>per capita</i> inicial			0,0002*** (0,0001)	0,0001** (0,0001)
Características do pai da criança				
Pai com emprego em tempo integral				-0,041** (0,019)
Anos de estudo				0,005*** (0,001)
Característica da mãe da criança				
Idade		-0,0002 (0,001)	-0,0002 (0,001)	-0,0001 (0,001)
Anos de estudo		0,008*** (0,001)	0,006*** (0,001)	0,003** (0,002)
Outros controles				
<i>Dummies</i> para RM		Sim	Sim	Sim
<i>Dummies</i> para os anos e meses da PME		Sim	Sim	Sim
Observações	13.748	13.748	13.748	12.838
Observação do tratamento	1.627	1.627	1.627	1.447
Pseudo- R^2	0,0006	0,034	0,035	0,036
Wald χ^2	10,21	559,21	565,35	545,99
Prob > χ^2	0,0014	0,0000	0,0000	0,0000

Fonte: PMEs de 1986 a 1995. Elaboração própria.

Notas: Erro-padrão robusto entre parênteses.

***significativo a 1%; **significativo a 5%; e *significativo a 10%.

Para se ter uma ideia da magnitude do efeito do ingresso materno no mercado de trabalho, optou-se por calcular também o efeito do tratamento sobre as crianças. Para tanto, a partir dos coeficientes dos modelos estimados, foi obtida a probabilidade de aprovação para cada criança da amostra. A partir daí, foi calculada uma média dessas probabilidades, restritas às crianças do grupo de tratamento. Em seguida, novamente apenas para as crianças do grupo de tratamento, foram obtidas estimativas da probabilidade de aprovação considerando que a *dummy* de tratamento fosse zero e uma média dessas novas probabilidades foi calculada. Dessa forma, o impacto da entrada da mãe no mercado sobre o desempenho escolar da criança foi obtido por meio da diferença entre essas duas médias, sendo o nível de significância obtido pela técnica *bootstrap*. Os resultados estão descritos na tabela 4. No último modelo, a probabilidade de aprovação das crianças tratadas é 66,9%; caso elas não fossem tratadas, a probabilidade seria de 68,5%. Embora exista uma diferença de 1,6 p.p., não podemos dizer que tal valor é estatisticamente diferente de zero.

TABELA 4

Efeito médio do tratamento sobre tratados: amostra de 1986 a 1995

Modelos	Probabilidade média de aprovação dos tratados	Probabilidade média de aprovação dos tratados caso não tivessem sido tratados	Estimativas
Modelo 1	0,657*** (0,011)	0,696*** (0,004)	-0,039*** (0,012)
Modelo 2	0,657*** (0,011)	0,683*** (0,005)	-0,026** (0,013)
Modelo 3	0,657*** (0,011)	0,679*** (0,005)	-0,022* (0,012)
Modelo 4	0,669*** (0,012)	0,685*** (0,005)	-0,016 (0,013)

Fonte: PMEs de 1986 a 1995. Elaboração própria.

Notas: Erro-padrão estimado via *bootstrap*.

***significativo a 1%; **significativo a 5%; e *significativo a 10%.

Procurando uma robustez maior aos resultados, foram utilizadas também técnicas de PSM. Optou-se por trabalhar com três estimadores: *kernel matching*, *nearest neighbor noreplacement* e cinco *neighbors*.¹⁷ Os resultados, apresentados na tabela A.1 do apêndice, corroboram quase todas as estimativas anteriores. Destacam-se, entretanto, as estimativas geradas por *kernel matching*, que foram estatisticamente significativas em todos os modelos analisados. Para o modelo 4, especificamente, a diferença da probabilidade de aprovação entre os grupos é 2 p.p. e o resultado é significativo a 10%.

17. Nas estimativas por *kernel matching*, optou-se por utilizar a especificação *epanechnikov kernel*.

Outra forma de testar a robustez das estimativas foi analisar os efeitos do trabalho materno para diferentes medidas de tratamento, de forma a identificar, separadamente, os efeitos da entrada da mãe no mercado de trabalho na condição de ocupação e na condição de desemprego.¹⁸ Para tanto, trabalhou-se com três medidas diferentes de tratamento, que serão explicitadas a seguir, utilizando sempre a composição de variáveis do modelo 4, que compreende, além dos demais controles, as variáveis referentes ao pai da criança, e por esta razão é o mais completo dos modelos.

A seguir são apresentados a composição dos tratamentos e os resultados das estimações a partir dos mesmos (tabela 5):

- tratamento 2: composto por uma *dummy* que atribui 1 às crianças cujas mães começaram a participar do mercado na condição de ocupadas e zero para aquelas cujas mães permaneceram inativas ou desempregadas;
- tratamento 3: formado por duas dummies que identificam se as mães começaram a trabalhar em tempo integral ou parcial,¹⁹ e mais uma vez atribuiu-se zero às crianças cujas mães permaneceram inativas ou desempregadas; e
- tratamento 4: constituído por uma variável contínua que representa a jornada, em horas semanais, das mães que começaram a trabalhar.

TABELA 5

Impacto do trabalho materno sobre a probabilidade de aprovação das crianças com idade entre 10 e 14 anos para diferentes tipos de tratamento: efeitos marginais dos modelos *logit* – amostra de 1986 a 1995

Variável dependente = “probabilidade de aprovação”	Tratamento 2	Tratamento 3	Tratamento 4
Crianças cujas mães começaram a participar do mercado na condição de empregadas	-0,020 (0,013)		
Crianças cujas mães começaram a trabalhar em tempo integral		-0,035** (0,014)	
Crianças cujas mães começaram a trabalhar em tempo parcial		0,038 (0,027)	
Jornada de trabalho da mãe			-0,0009*** (0,0004)

Fonte: PMEs de 1986 a 1995. Elaboração própria.

Notas: Erro-padrão robusto entre parênteses.

***significativo a 1%; **significativo a 5%; e *significativo a 10%.

18. Além dessas medidas de tratamento, estimou-se também um modelo que contempla, além dos controles já utilizados, a variável que capta o salário médio da mãe. A ideia era separar o efeito renda do efeito tempo da mãe. Entretanto, as estimativas não foram significativas.

19. A jornada é dita parcial quando são trabalhadas até 20 horas semanais. Já as jornadas de mais de 20 horas são consideradas de tempo integral.

Iniciando a análise pelo tratamento 2, observa-se que o efeito causado pelo ingresso da mãe no mercado de trabalho não foi estatisticamente significativo. Já as estimativas dos tratamentos 3 e 4 corroboram os resultados da literatura, que atribuem efeitos mais restritivos do emprego materno às maiores jornadas de trabalho. Nota-se, por exemplo, que o fato de a mãe trabalhar mais de 20 horas semanais, ou seja, em jornada integral, reduz as chances de aprovação das crianças, em média, em 3,5 p.p., lembrando que essa estimativa é relativa às crianças cujas mães permaneceram na inatividade ou desempregadas.

Em uma última abordagem, os efeitos da entrada da mãe no mercado de trabalho foram analisados de forma mais detalhada, de acordo com a escolaridade da mãe da criança.²⁰ Para isso, trabalhamos com a interação da variável tratamento e o grupo de escolaridade da mãe.²¹

Os resultados apresentados na tabela 6 representam, mais uma vez, estimativas do modelo 4. Optou-se, nesse caso, por apresentar somente os resultados referentes ao tratamento 1, ou seja, aquele que identifica se as crianças tiveram suas mães ingressando no mercado de trabalho. Como pode ser observado, somente a estimativa da diferença na probabilidade de aprovação entre crianças tratadas e não tratadas, para filhos de mães mais escolarizadas, foi significativa. Para essas crianças, o fato de pertencer ao grupo de tratamento reduz as chances de aprovação em 3,2 p.p.

TABELA 6

Efeito do tratamento sobre tratados para diferentes grupos de escolaridade materna: amostra de 1986 a 1995

Amostras	Estimativas	Número de observações do tratamento	Número de observações totais
Crianças com mães de até quatro anos de estudo	0,009	926	8.165
Crianças com mães de mais de quatro anos de estudo	-0,032**	522	4.689

Fonte: PMEs de 1986 a 1995. Elaboração própria.

Nota: ***significativo a 1%; **significativo a 5%; e *significativo a 10%.

Em suma, os resultados mostraram que, em relação à subamostra da PME referente ao período 1986-1995, não se observou impacto, no modelo mais completo, do ingresso materno no mercado de trabalho sobre a probabilidade de aprovação

20. Também foi estimado um modelo interagindo a idade da criança com a *dummy* de tratamento, a fim de verificar se o efeito da presença da mãe varia conforme a idade da criança. Entretanto, devido à não significância estatística da interação, optou-se por não apresentá-lo.

21. Para tanto, foi utilizado o comando *inteff* do software STATA, que trata de efeitos de interações e respectivos erros-padrão em modelos *logit*. Nesse caso, estimou-se o modelo 4 incluindo uma *dummy* para o grupo de escolaridade (variável binária igual a um quando a mãe apresenta escolaridade maior que a mediana e igual a zero, caso contrário) e a interação dessa *dummy* com a variável *tratamento*. As estimativas apresentadas na tabela 6 representam a diferença na probabilidade de aprovação, para crianças cujas mães são menos escolarizadas, entre crianças tratadas e não tratadas (efeito marginal da *dummy* de tratamento) e a diferença na probabilidade de aprovação, para crianças cujas mães são mais escolarizadas, entre crianças tratadas e não tratadas (efeito marginal da *dummy* de tratamento + efeito marginal da interação).

escolar da criança. Todavia, é importante ressaltar que houve um aumento expressivo desse efeito ao considerar a jornada de trabalho da mãe: crianças cujas mães começaram a trabalhar em tempo integral tiveram sua probabilidade de aprovação reduzida, em média, em 3,5 p.p. E, com relação à escolaridade materna, crianças cujas mães são mais escolarizadas e que ingressaram no mercado de trabalho têm menores chances de concluir com êxito o ano letivo do que os filhos de mães mais escolarizadas que permaneceram inativas.

4.2 Características descritivas e análise dos modelos *logit* para a amostra de 2002 a 2006

As características da amostra obtida a partir da PME nova estão apresentadas na tabela 7. Mais uma vez, as médias das variáveis seguem acompanhadas dos respectivos erros-padrão e dos testes de diferença de médias.

Em relação às características individuais da criança, não se observaram diferenças significativas para as variáveis de *idade*, *sexo*, *trabalho infantil* e *atraso escolar* entre as crianças do grupo de tratamento e controle. Em média, as crianças analisadas têm 12 anos de idade e meio ano de atraso escolar; quanto ao gênero, 50% são meninos e em torno de 1% participa do mercado de trabalho.

As características familiares “estrutura familiar” e “renda familiar *per capita*” mostraram-se estatisticamente diferentes entre os grupos. Em média, em relação ao grupo de tratamento, essas variáveis apresentaram valores inferiores aos do grupo de controle. Essas variáveis levam a mãe para o mercado de trabalho e também devem influenciar o desempenho educacional das crianças; sendo assim, devem ser controladas. Por outro lado, observa-se que os grupos não diferiram quanto ao número de irmãos. E, em relação à variação proporcional da renda familiar *per capita*, nota-se que enquanto as famílias do grupo de controle tiveram uma variação anual positiva de 13,3%, para o grupo de tratamento o aumento representou 17,8% do valor da renda inicial. Vale mencionar que, diferentemente do observado para a amostra anterior, a variação anual de renda se mostrou positiva, indicando que, para essa amostra, as famílias tiveram um ganho de renda.

Quanto às características maternas, a educação da mãe não apresentou divergências entre os grupos. No entanto, a variável idade e a relativa ao passado ocupacional da mãe mostraram diferenças consideráveis, já que a proporção de mães, pertencentes ao grupo de tratamento, que tinham participado do mercado de trabalho no ano anterior à primeira entrevista é seis vezes superior ao do grupo de controle. Isto pode estar indicando uma preferência maior pelo trabalho entre as mães do grupo de tratamento.

Em relação às características dos pais das crianças, a proporção daqueles que trabalhavam em tempo integral é superior no grupo de controle, assim como é maior o nível de escolaridade dos pais para esse grupo.

TABELA 7

Características descritivas dos grupos: média e desvio-padrão – amostra de 2002 a 2006

Variáveis	Grupo de controle		Grupo de tratamento		Diferenças significativas para nível de confiança de 95%
	Média	Desvio-padrão	Média	Desvio-padrão	
Idade da criança	12,02	1,37	12,07	1,40	
Sexo masculino	51%	0,50	49%	0,50	
Trabalho infantil	1%	0,10	1%	0,10	
Atraso escolar	0,48	0,93	0,50	0,90	
Número de irmãos de 0 a 14 anos	1,08	1,11	1,22	1,17	
Número de irmãos de 15 anos ou mais	0,46	0,76	0,38	0,69	
Família completa	93%	0,25	80%	0,40	*
Renda familiar inicial <i>per capita</i>	216,20	340,38	114,98	140,77	*
Δ da renda familiar <i>per capita</i>	28,85	264,72	20,44	133,99	
% de pais trabalhando em tempo integral	92%	0,27	78%	0,42	*
Educação do pai	7,58	4,04	7,04	3,78	
Idade da mãe	39,36	7,61	37,83	6,69	*
Educação da mãe	7,12	3,78	6,67	3,47	
% de mães que estavam no mercado de trabalho no ano anterior	4%	0,005	24%	0,03	*
Residência em Salvador	7%	0,26	8%	0,27	
Residência em Recife	10%	0,30	13%	0,34	
Residência em Belo Horizonte	16%	0,37	27%	0,45	*
Residência no Rio de Janeiro	29%	0,46	18%	0,39	*
Residência em São Paulo	22%	0,41	21%	0,41	
Residência em Porto Alegre	15%	0,36	13%	0,33	
Número de observações	1.187		211		

Fonte: PMEs de 2002 a 2006. Elaboração própria.

Nota: *indica diferença significativa.

Por último, com relação à região de residência das crianças, cabe ressaltar que as diferenças são significativas somente para as RMs do Rio de Janeiro e Belo Horizonte.

Em suma, embora para a maior parte das variáveis, os testes de diferença de média não tenham apontado divergências entre os grupos, esta análise preliminar mostrou diferenças importantes para o grupo de variáveis que caracterizam as famílias e também para a variável referente ao passado ocupacional da mãe. Estes resultados atestam, portanto, a relevância de incluir tais controles nos exercícios de regressão.

Para estimar a probabilidade de aprovação da criança foram construídos cinco modelos *logit*. Os quatro primeiros têm o mesmo desenho da análise anterior; e no último modelo foi incluída a variável representativa do passado ocupacional da mãe. A ideia é utilizar essa última variável para testar a robustez dos resultados.

A tabela 8 apresenta os efeitos marginais, referentes aos modelos mencionados anteriormente, calculados para indivíduos de características médias da amostra. Os resultados das regressões mostram que o fato de uma criança fazer parte do grupo de tratamento tem efeitos restritivos sobre a probabilidade de aprovação escolar. Considerando os cinco modelos estimados, o efeito da *dummy* de tratamento varia entre $-5,9$ p.p. e $-7,1$ p.p., sendo estatisticamente significativo, a pelo menos 10%, em todos os modelos.

Quanto às demais variáveis, é possível observar que a maioria delas não se mostrou estatisticamente significativa. Provavelmente esses efeitos estão sendo captados por outras variáveis. Por exemplo, o atraso escolar é uma variável forte que capta todo um histórico educacional da criança. Ao excluí-la dos modelos, variáveis como número de irmãos, idade da criança e educação da mãe passam a ser estatisticamente significativas. Esta última variável também perde significância quando controlada a “renda familiar inicial *per capita*”. Em relação à variável atraso escolar, é possível notar que, assim como esperado, quanto maior o atraso escolar da criança, ou seja, quanto mais distante ela estiver da escolaridade ideal, menores são as chances de aprovação. Em média, o aumento de um ano na defasagem reduz a probabilidade de aprovação em 2,3 p.p., resultado análogo ao anterior.

Com relação aos modelos 4 e 5 – que incluem características paternas da criança – pode-se observar um efeito restritivo da *dummy* que identifica pais com emprego em tempo integral.²² Para essas crianças, a probabilidade de aprovação fica reduzida em 6,8 p.p. Esse resultado sugere que, assim como a presença materna, a presença do pai no domicílio é um evento bastante relevante para os bons resultados escolares da criança. Lembrando que, devido à renda estar controlada, observa-se o efeito puro do tempo despendido pelo pai em casa.

Por fim, como mostrado anteriormente, os grupos de tratamento e controle apresentaram grandes divergências quanto à variável referente ao passado ocupacional da mãe, sendo relevante sua incorporação como variável de controle nos exercícios de regressão. Os resultados mostram que mães que participaram do mercado de trabalho no ano anterior tiveram efeitos positivos sobre a probabilidade de aprovação das crianças. Ou seja, controlada a defasagem idade-série, o fato de a mãe ter participado do mercado anteriormente, impactar positivamente a probabilidade de aprovação da criança sinaliza que essas crianças já eram melhores em termos de desempenho escolar.

22. Para a *dummy* que identifica pais com emprego em tempo integral, optou-se por apresentar também as probabilidades previstas nos pontos. Considerando o modelo 5, a probabilidade prevista de aprovação é de 0,90 para as crianças cujos pais trabalham em tempo integral e 0,97 para os filhos cujos pais não trabalham em tempo integral.

TABELA 8

**Impacto do trabalho materno sobre a probabilidade de aprovação das crianças:
efeitos marginais dos modelos *logit* – amostra de 2002 a 2006**

Variável dependente = "probabilidade de aprovação"	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Crianças cujas mães participaram do mercado	-0,064** (0,030)	-0,064** (0,032)	-0,059* (0,031)	-0,060* (0,035)	-0,071* (0,037)
Características das crianças					
Trabalho infantil		0,036 (0,056)	0,035 (0,055)	0,033 (0,052)	0,035 (0,050)
Anos de atraso escolar		-0,023*** (0,008)	-0,023*** (0,008)	-0,023*** (0,008)	-0,024*** (0,008)
Sexo masculino		-0,013 (0,018)	-0,014 (0,017)	-0,023 (0,018)	-0,022 (0,018)
Idade		-0,007 (0,007)	-0,006 (0,006)	-0,008 (0,007)	-0,007 (0,007)
Características da família da criança					
Família completa		0,006 (0,027)	-0,005 (0,025)		
Número de irmãos com idade <= 14		-0,012 (0,009)	-0,011 (0,008)	-0,008 (0,009)	-0,007 (0,009)
Número de irmãos com idade > 15		0,010 (0,014)	0,009 (0,014)	0,017 (0,016)	0,016 (0,016)
Varição da renda familiar <i>per capita</i>			0,0001 (0,0001)	0,0001 (0,0001)	0,0001 (0,0001)
Renda familiar <i>per capita</i> inicial			0,00006** (0,00003)	0,00007** (0,00003)	0,0001** (0,00003)
Características do pai da criança					
Pai com emprego em tempo integral				-0,068** (0,029)	-0,069** (0,028)
Anos de estudo				-0,0001 (0,003)	-0,0001 (0,003)
Característica da mãe da criança					
Idade		-0,002 (0,002)	-0,002 (0,001)	-0,001 (0,002)	-0,001 (0,002)
Anos de estudo		0,005* (0,003)	0,003 (0,003)	0,003 (0,003)	0,003 (0,003)
Mãe que participou do mercado no ano anterior					0,045* (0,027)
Outros controles					
<i>Dummies</i> para RM		Sim	Sim	Sim	Sim
<i>Dummies</i> para os anos e meses da PME		Sim	Sim	Sim	Sim
Observações	1.398	1.398	1.398	1.265	1.265
Observação do tratamento	211	211	211	165	165
Pseudo- R^2	0,006	0,068	0,070	0,080	0,082
Wald χ^2	5,63	72,23	76,21	77,27	78,65
Prob > χ^2	0,018	0,000	0,000	0,000	0,000

Fonte: PMEs de 2002 a 2006. Elaboração própria.

Notas: Erro-padrão robusto entre parênteses.

***significativo a 1%; **significativo a 5%; e *significativo a 10%.

A fim de se ter uma ideia da magnitude desse efeito, repetiu-se o procedimento da estimação anterior, ou seja, levou-se em conta somente o grupo de crianças tratadas. Considerando-se como exemplo o modelo 5, pode-se concluir que crianças tratadas teriam sua probabilidade de aprovação aumentada de 82% para 89%, caso não fossem tratadas. Ou seja, o impacto do trabalho materno sobre a probabilidade de aprovação da criança é negativo e tem uma magnitude de 7 p.p. – um efeito bem elevado. Se olharmos pelo ponto de vista da reprovação, os resultados chamam ainda mais a atenção: observa-se que a taxa de reprovação é maior, em mais de 60%, para o grupo de tratamento, já que a probabilidade média de reprovação dos tratados é de 18% e, no caso de o grupo não ser tratado, a taxa cai para 11%. A tabela 9 apresenta o efeito do tratamento sobre tratados para os cinco modelos estimados. Um ponto interessante é o aumento da probabilidade de aprovação nos dados dos anos 2000, relativamente aos dados dos anos 1980 e 1990, que deve estar refletindo efeitos de políticas educacionais de correção de fluxo que tem ocorrido nos últimos anos.

TABELA 9

Efeito do tratamento sobre tratados: amostra de 2002 a 2006

Modelos	Probabilidade média de aprovação dos tratados	Probabilidade média de aprovação dos tratados caso não tivessem sido tratados	Estimativas
Modelo 1	0,826*** (0,028)	0,890*** (0,009)	-0,064** (0,030)
Modelo 2	0,807*** (0,030)	0,880*** (0,013)	-0,073** (0,034)
Modelo 3	0,808*** (0,030)	0,877*** (0,013)	-0,069** (0,035)
Modelo 4	0,817*** (0,035)	0,886*** (0,013)	-0,069* (0,039)
Modelo 5	0,818*** (0,033)	0,894*** (0,014)	-0,076** (0,037)

Fonte: PMEs de 2002 a 2006. Elaboração própria.

Notas: Erro-padrão estimado via *bootstrap*.

***significativo a 1%; **significativo a 5%; e *significativo a 10%.

Também foram utilizadas técnicas de PSM. Os resultados, mostrados na tabela A.2 do apêndice, sustentam as estimativas apresentadas anteriormente. Uma única exceção fica por conta da estimativa gerada por cinco *neighbors* que, apesar de conter magnitude semelhante, não se mostrou estatisticamente significativa para o modelo 4.

Seguindo com a análise, a mesma estratégia de estimar diferentes medidas de tratamento foi adotada para essa amostra. Para tanto, trabalhou-se com quatro tipos

diferentes de tratamento, utilizando sempre o modelo 5, ou seja, o mais completo dentre os modelos.²³ Os resultados estão apresentados na tabela 10.

TABELA 10

Impacto do trabalho materno sobre a probabilidade de aprovação das crianças com idade entre 10 e 14 anos para diferentes tipos de tratamento: efeitos marginais dos modelos *logit* – amostra de 2002 a 2006

Variável dependente = "probabilidade de aprovação"	Tratamento 2	Tratamento 3	Tratamento 4
Crianças cujas mães começaram a participar do mercado na condição de empregadas	-0,058 (0,040)		
Crianças cujas mães começaram a trabalhar em tempo integral		-0,013 (0,036)	
Crianças cujas mães começaram a trabalhar em tempo parcial		-0,145* (0,087)	
Jornada de trabalho da mãe			-0,0003 (0,0007)
Número de observações	Total = 1.265 Tratamento = 139	Total = 1.265 Tempo integral = 95 Tempo parcial = 44	Total = 1.265 Tratamento = 139

Fonte: PMEs de 2002 a 2006. Elaboração própria.

Notas: Erro-padrão robusto entre parênteses.

***significativo a 1%; **significativo a 5%; e *significativo a 10%.

Iniciando a análise pelo tratamento 2, observa-se, apesar de apresentar sinal negativo, um efeito não significativo do ingresso da mãe no mercado de trabalho na condição de emprego.

As estimativas dos tratamentos 3 e 4, desta vez, não corroboram os resultados da literatura que atribuem efeitos mais restritivos do emprego materno às maiores jornadas de trabalho. Chama a atenção a magnitude da *dummy* que capta o efeito do trabalho materno em tempo parcial (-14,4 p.p.). Esse resultado, provavelmente, reflete o reduzido montante de observações que compõe esse grupo de tratamento. Mais especificamente, dentre as 165 crianças cujas mães deixaram de ser inativas, apenas 44 passaram a trabalhar em tempo parcial. Por outro lado, analisando-se as características descritivas deste grupo de mulheres, observa-se que a porcentagem destas que trabalharam no ano anterior é relativamente menor quando comparada à média observada para o grupo de tratamento (20% versus 24%). Ou seja, uma possibilidade é que essas mulheres que passam a trabalhar em tempo parcial são as mais presentes em casa e, então, a sua ausência acaba por fazer uma diferença maior.

23. Também foi estimado o modelo 5 acrescido da variável que capta o salário médio da mãe. Novamente, não se conseguiu separar o efeito renda do efeito tempo da mãe. Mais uma vez os resultados não foram significativos.

Assim como foi realizado anteriormente, decidiu-se analisar os efeitos da entrada da mãe no mercado de trabalho utilizando um modelo que contempla uma interação da variável de tratamento com uma *dummy* que identifica mães com nível de escolaridade maior que a mediana. Como pode ser observado, somente a estimativa da diferença na probabilidade de aprovação entre crianças tratadas e não tratadas, para filhos de mães mais escolarizadas, foi significativa. Para essas crianças, o fato de pertencer ao grupo de tratamento reduz suas chances de aprovação em 17,6 p.p. Ou seja, o efeito médio do trabalho materno é significativamente mais custoso para os filhos de mães mais escolarizadas. Embora este resultado seja condizente com os encontrados na literatura – a ideia é que em função do maior capital humano, a ausência das mães mais educadas acaba sendo mais sentida pelas crianças – novamente aqui este efeito é bastante elevado por ser devido ao número restrito de observações para este grupo.

TABELA 11

Efeito do tratamento sobre tratados para diferentes grupos de escolaridade materna: amostra de 2002 a 2006

Amostras	Estimativas	Número de observações do tratamento	Número de observações totais
Crianças com mães de até sete anos de estudo	0,003	83	569
Crianças com mães de mais de sete anos de estudo	-0,176***	82	696

Fonte: PMEs de 2002 a 2006. Elaboração própria.

Nota: ***significativo a 1%; **significativo a 5%; e *significativo a 10%.

Em suma, os resultados mostraram que, para a amostra referente ao período de 2002 a 2006, o impacto do trabalho materno apresentou uma magnitude expressiva, com valores bem superiores aos encontrados para a amostra de 1986 a 1995. O que pode explicar estas diferenças de resultados?

Primeiramente, as próprias mudanças metodológicas da PME podem estar explicando estes resultados. A PME nova capta com mais rigor a participação no mercado de trabalho dos indivíduos. São dois pontos principais que impactam neste trabalho. O período de referência para identificar a procura por trabalho aumenta de sete para 30 dias. Assim, indivíduos que eram classificados como inativos agora são classificados como desempregados. Na PME antiga o trabalho não remunerado só era considerado ocupação se fosse acima de 15 horas por semana. Na nova PME, basta ultrapassar 1 hora. Dessa forma, indivíduos que eram classificados como inativos agora são classificados como ocupados. Nos dois casos, pessoas do grupo de tratamento atual antes estavam no grupo de controle. Assim, uma possibilidade é que na PME antiga o impacto seja subestimado, visto que havia mulheres tratadas no grupo de controle.

Um segundo ponto a ser considerado é com relação às diferenças observadas na variável “variação anual da renda *per capita* da família” entre os períodos. Observa-se que a variação anual da renda passou de negativa na amostra de 1986-1995 para positiva no período 2002-2006. A partir daí, pode-se pensar que, para as crianças

da primeira amostra, o efeito positivo da renda proveniente do trabalho materno se mostrou mais importante para o desempenho escolar da criança relativamente ao efeito negativo da ausência da mãe no domicílio. Em contrapartida, as crianças da segunda amostra tiveram uma variação anual positiva na renda familiar. E, para esse grupo, o efeito da renda extra advinda do trabalho da mãe pode não ter sido tão relevante, de forma que o impacto negativo da ausência da mãe se sobressaiu.

Outro fato que merece ser destacado é a expressiva diferença entre as taxas de reprovação nos dois períodos analisados. Enquanto na PME antiga as taxas de reprovação ficam em torno de 35%, na PME nova a taxa é de apenas 19%. Dessa forma, pode-se dizer que a reprovação, no período 1986-1995, era considerada como a regra. Já no período mais recente a regra passou a ser a aprovação das crianças e, por isso, os efeitos negativos da ausência da mãe passaram a ser captados.

E, por fim, mas não menos importante, a diferença de resultados pode estar refletindo uma mudança, em termos de escolaridade materna, entre os períodos analisados. Como observado, as características descritivas das amostras indicam que as mães do período mais antigo (1986-1995) que começaram a participar do mercado de trabalho tinham, em média, 3,8 anos de estudo. Já na amostra mais recente (2002-2006), essa média chega a 6,7 anos de escolaridade – um aumento expressivo de 76,3%. Assim como mencionado anteriormente, a literatura aponta efeitos restritivos mais intensos do trabalho materno conforme aumenta a escolaridade da mãe da criança. Dessa forma, pode-se dizer que as crianças provenientes da amostra da PME nova apresentaram uma sensibilidade maior à ausência da mãe no domicílio não somente pela simples redução no tempo de dedicação materna aos filhos, mas também em função da ausência de uma mãe mais escolarizada. Para os dois subperíodos, o impacto do tratamento é mais negativo para as crianças filhas de mães mais escolarizadas. Como este grupo é mais representativo na PME mais nova, os resultados aparecem.

5 CONSIDERAÇÕES FINAIS

O objetivo deste estudo foi analisar o impacto do trabalho materno sobre a probabilidade de aprovação das crianças de 10 a 14 anos de idade. Para tanto foram utilizadas duas subamostras da PME referentes aos períodos 1986-1995 e 2002-2006. Tal decisão se apoiou nas vantagens exclusivas do uso de cada uma delas: enquanto a amostra referente à metodologia antiga ofereceu uma quantidade maior de graus de liberdade e ainda estava livre dos problemas relacionados ao sistema de progressão continuada, ela teve a desvantagem de não ter uma variável que desempenhasse a função de controlar o passado ocupacional da mãe, variável que só pôde ser incorporada na amostra mais recente.

A contribuição deste trabalho, além de analisar um atributo da família pouco explorado nos estudos brasileiros sobre o tema, está em utilizar os dados da PME, os quais proporcionaram um trabalho diferenciado em virtude de se poder trabalhar com um painel de indivíduos.

Os resultados mostraram que, para a amostra mais antiga, o fato de a mãe participar do mercado de trabalho não teve impacto sobre a probabilidade de aprovação escolar da criança. Todavia, houve um aumento expressivo desse efeito ao se considerar a jornada de trabalho e a escolaridade da mãe: crianças cujas mães começaram a trabalhar em tempo integral tiveram sua probabilidade de aprovação reduzida, em média, em 3,5 p.p.; para as crianças, filhas de mães mais escolarizadas, o impacto negativo é 3,2 p.p.

Já na amostra mais recente, os resultados foram mais expressivos: o impacto do trabalho materno apresentou uma magnitude de 7,1 p.p. Neste caso, o efeito mais negativo para a extensão da jornada de trabalho não aparece, enquanto o impacto mais negativo observado para as crianças filhas de mães mais escolarizadas permanece.

Essa divergência de resultados pode ser decorrente: *i*) das mudanças metodológicas da PME; *ii*) das diferenças da variável “variação anual da renda *per capita* da família”; *iii*) das mudanças das “regras” relativas à aprovação/reprovação nas escolas; ou, por fim, *iv*) do aumento da escolarização das mães entre os períodos analisados. Estudos futuros podem buscar a confirmação da importância destes elementos e, em caso positivo, qual a contribuição de cada um deles no resultado final.

ABSTRACT

The purpose of this paper is to analyze the impact of maternal labor on the probability of school progression for Brazilian children with age between 10 and 14 years. Two sub-samples from the data set of the Monthly Employment Survey (Pesquisa Mensal de Emprego – PME), referring to periods from 1986 to 1995 and from 2002 to 2006 were used. The results show that, for the oldest sample, maternal ingression in the labor market does not influence the child's school progression probability. However, from 2002 to 2006, the results are more expressive: maternal labor impacts school progression probability negatively in 7.1 percentage points. This divergent behavior may be attributed to: *i*) changes in PME methodology; *ii*) difference in sub-samples with respect to the per capita income variable; *iii*) changes in school progression/repetition “rules”; or, finally, *iv*) increase in maternal education levels between the analyzed periods. Future studies may identify the importance of these elements and their contribution to our final results

REFERÊNCIAS

- ALBERNAZ, A.; FERREIRA, F.; FRANCO, C. Qualidade e equidade na educação fundamental brasileira. **Pesquisa e Planejamento Econômico**, Rio de Janeiro: Ipea, v. 33, n. 3, p. 453-476, 2002.
- BARROS, R. P. de. *et al.* **Determinantes do desempenho educacional no Brasil**. Rio de Janeiro: Ipea, 2001 (Texto para Discussão, n. 834).
- BLAU, F. D.; GROSSBERG, A. J. Maternal labor supply and children's cognitive development. **The Review of Economics and Statistics**, v. 74, n. 3, p. 474-481, 1992.
- CAVALIERE, C. H. Trabalho infantil e desempenho escolar. In: ENCONTRO NACIONAL DE ECONOMIA, 28., 2000, Campinas. **Anais....** Campinas: ANPEC, 2000.

CURI, A. Z.; MENEZES-FILHO, N. A. O mercado de trabalho brasileiro é segmentado? Alterações no perfil da informalidade e nos diferenciais de salários nas décadas de 1980 e 1990. **Estudos Econômicos**, São Paulo, v. 36, n. 4, p. 867-899, out./dez. 2006.

FELICIO, F.; FERNANDES, R. O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de São Paulo. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO EM ECONOMIA, 2005, Natal. **Anais...** Natal: ANPEC, 2005.

FERNANDES, R.; CHAHAD, J. P. Z. Unemployment insurance and transitions in the labor market: an evaluation of the Brazilian program. **Brazilian Review of Econometrics**, v. 22, 2002.

_____.; FELÍCIO, F. The entry of the wife into the labor force in response to the husband's unemployment: a study of the added worker effect in Brazilian metropolitan areas. **Economic Development and Cultural Change**, v. 53, p. 887-911, July 2005.

GONZAGA, G. M.; MENEZES-FILHO, N. A.; CAMARGO, J. M. Os efeitos da redução da jornada de trabalho de 48 para 44 horas semanais em 1988. **Revista Brasileira de Economia**, v. 57, n. 2, p. 369-400, 2003.

GREGG, P. *et al.* The effects of a mother's return to work decision on child development in the UK. **The Economic Journal**, v. 115, p. 48-80, 2005.

HANUSHECK, E. The failure of input-based schooling policies. **Economic Journal**, v. 113, n. 485, p. 64-98, 2003.

_____. **Education production functions: developed country evidence.** International Encyclopedia of Education. 3. ed. Nov. 2007. Disponível em: <<http://edpro.stanford.edu/hanushek/admin/pages/files/uploads/hanushek%202010%20prod%20functions%20international%20encyclopedia.pdf>>

HAVEMAN, R.; WOLFE, B. The determinants of children's attainments: a review of methods and findings. **Journal of Economic Literature**, v. 33, n. 4, p. 1.829-1.878, 1995.

_____.; _____.; SPAULDING, J. Childhood events and circumstances influencing high school completion. **Demography**, v. 28, n. 1, p. 133-157, 1991.

HOROWITZ, A. W.; SOUZA, A. P. **The dispersion of intra-household human capital across children:** a measurement strategy and evidence. Department of Economics, Vanderbilt University, 2004 (Working Papers, n. 0408). Disponível em: <<http://ideas.repec.org/p/van/wpaper/0408.html>>

KASSOUF, A. L. **O trabalho infantil no Brasil.** São Paulo, Piracicaba: Deas-ESALQ-USP, 1999. 110 p. (Tese de livre docência).

KREIN, S. F.; BELLER, A. H. Educational attainment of children from single-parent families: differences by exposure, gender, and race. **Demography**, v. 25, n. 2, p. 221-234, 1988.

LEON, F. L.; MENEZES-FILHO, N. A. Reprovação, avanço e evasão escolar no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro: Ipea, v. 32, n. 3, 2002.

MACHADO, A. F.; RIBAS, R. P.; PENIDO, M. Mobilidade entre estados de pobreza e inserção no mercado de trabalho: uma análise para o Brasil metropolitano em 2004. **Economia Aplicada**, v. 11, n. 2, p. 253-279, 2007.

MACHADO, D. C. **Escolaridade das crianças no Brasil:** três ensaios sobre a defasagem idade-série. 2005. 142f. Tese (Doutorado) – Departamento de Economia da PUC-Rio, Rio de Janeiro, 2005.

_____.; GONZAGA, G. O impacto dos fatores familiares sobre a defasagem idade-série de crianças no Brasil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 61, n. 4, p. 449-476, 2007.

MENEZES-FILHO, N. A.; PICCHETTI, P. Os determinantes da duração do desemprego em São Paulo. **Pesquisa e Planejamento Econômico**, Rio de Janeiro: Ipea, v. 30, n. 1, 2000.

MILNE, A. M. *et al.* Single parents, working mothers, and the educational achievement of school children. **Sociology of Education**, East Lansing, v. 59, p. 125-139, 1986.

RIBAS, R.; SOARES, S. **Sobre o painel da Pesquisa Mensal de Emprego (PME) do IBGE**. Rio de Janeiro: Ipea, 2008 (Texto para Discussão, n. 1.348).

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, v. 70, p. 41-55, 1983.

RUHM, C. J. **Parental employment and child cognitive development**. Greensboro: University of North Carolina, 2002.

SILVA, N. V.; HASENBALG, C. Tendências da desigualdade educacional no Brasil. **Dados**, v. 43, n. 3, 2000.

WALDFOGEL, J. *et al.* The effects of early maternal employment on child cognitive development. **Demography**, v. 39, n. 2, p. 369-392, 2002.

(Originais submetidos em setembro de 2008. Última versão recebida em outubro de 2010. Aprovada em dezembro de 2010.)

APÊNDICE

TABELA A.1

PSM: amostra de 1986 a 1995

Modelos	<i>Kernel matching</i>	<i>Nearest neighbor</i>	<i>5 neighbors</i>
Modelo 2	-0,029** (0,013)	-0,039** (0,018)	-0,032** (0,016)
Modelo 3	-0,026** (0,013)	-0,029* (0,017)	-0,020 (0,016)
Modelo 4	-0,0197* (0,011)	-0,006 (0,013)	-0,0164 (0,015)

Fonte: PMEs de 1986 a 1995. Elaboração própria.

Notas: Erro-padrão estimado via *bootstrap*.

***significativo a 1%; **significativo a 5%; e *significativo a 10%.

TABELA A.2

PSM: amostra de 2002 a 2006

Modelos	<i>Kernel matching</i>	<i>Nearest neighbor</i>	<i>5 neighbors</i>
Modelo 2	-0,069** (0,030)	-0,078** (0,037)	-0,065* (0,035)
Modelo 3	-0,066** (0,030)	-0,077** (0,037)	-0,063* (0,035)
Modelo 4	-0,062* (0,032)	-0,067* (0,040)	-0,059 (0,038)
Modelo 5	-0,0603* (0,034)	-0,067* (0,040)	-0,071* (0,039)

Fonte: PMEs de 2002 a 2006. Elaboração própria.

Notas: Erro-padrão estimado via *bootstrap*.

***significativo a 1%; **significativo a 5%; e *significativo a 10%.

GASTO EM P&D E PODER DE MERCADO: TEORIA E EVIDÊNCIA PARA O BRASIL*

Eduardo Correia**

Rodrigo Moita**

Desde Schumpeter (1950), até trabalhos mais recentes como Aghion *et al.* (2002), existe uma ampla literatura devotada à análise da relação entre inovação e poder de mercado. Este artigo segue essa tradição e analisa a relação entre gasto em pesquisa e desenvolvimento (P&D) e poder de mercado na indústria brasileira. Usamos uma base de dados nova: a pesquisa IBMEC-Sensus 2008. Essa base tem como principal diferencial o fato de incorporar em seu questionário uma estimativa da elasticidade da demanda da firma, variável que usamos como instrumento exógeno para o poder de mercado. Usamos o modelo de Aghion *et al.* (2002), que propõe a existência de uma relação na forma de U-invertido entre P&D e poder de mercado – medido pelo índice de Lerner (IL) – para fornecer uma justificativa teórica para a especificação econométrica adotada. Os resultados obtidos não rejeitam a hipótese de uma relação entre P&D e poder de mercado na forma de U-invertido. Porém, quando estimamos o modelo usando o inverso da elasticidade da demanda como um instrumento para IL, essa relação perde significância, indicando que a relação entre P&D e poder de mercado pode simplesmente ser fruto da determinação simultânea das duas variáveis, não havendo uma relação de causalidade entre elas.

1 INTRODUÇÃO

Este artigo procura analisar a relação entre gasto em pesquisa e desenvolvimento (P&D) e poder de mercado na indústria brasileira. O debate sobre como a estrutura do mercado afeta o incentivo à inovação vem de longa data, tendo provavelmente como marco inicial o trabalho de Schumpeter (1950), que afirmava que empresas maiores, e com maior poder de mercado, gastariam proporcionalmente mais em P&D. Desde então uma extensa literatura sobre o tema se desenvolveu, com destaque para os trabalhos pioneiros de Scherer (1967), Kamien e Schwartz (1982), Cohen, Levin e Mowery (1987), e, mais recentemente, Aghion *et al.* (2002). O presente trabalho pretende contribuir para esse debate usando uma base de dados da indústria brasileira – a pesquisa IBMEC-Sensus 2008.

Do ponto de vista das firmas, a inovação pode ser responsável pela competitividade e pelo sucesso ou fracasso no ambiente competitivo do mercado. Do ponto de vista da sociedade, a inovação feita pelas firmas é responsável por

* Somos gratos pelos comentários e sugestões do editor da PPE e de dois pareceristas anônimos. Agradecemos também pela oportunidade de expor e debater nosso *paper* no Seminário Brown Bag do Instituto de Ensino e Pesquisa (Insper) e no XXXII Encontro Brasileiro de Econometria.

** Professor do Instituto de Ensino e Pesquisa (Insper).

parte considerável do avanço tecnológico e do crescimento econômico. O efeito da inovação pode transbordar para além dos limites da firma inovadora, pois a inovação feita por uma firma pode ser assimilada (ou comprada, ou imitada) por outras firmas, que também se apropriarão dos benefícios da inovação. Em suma, o processo inovador gera uma externalidade positiva.

Reconhecendo esta importância, vários países adotam algum tipo de incentivo à atividade inovadora. O Brasil não foge à regra, e adota uma série de medidas que procuram incentivar e subsidiar esta atividade. A chamada Lei do Bem (Lei nº 11.196/2005) ampliou os benefícios existentes e instituiu outros novos para estimular a inovação. Estes benefícios vão desde os fiscais para a atividade inovadora, isenção tributária de insumos estratégicos, linhas de financiamento – Banco Nacional de Desenvolvimento Econômico e Social (BNDES) e da Financiadora de Estudos e Projetos (FINEP) –, formação de recursos humanos etc. Enfim, o Brasil gasta recursos e faz um grande esforço, comparado a países com renda *per capita* semelhante, para incentivar a inovação tecnológica¹ (CGEE; ANPEI, 2009).

No entanto, vemos uma grande heterogeneidade entre setores da economia brasileira quanto ao investimento em inovação. Por exemplo, as indústrias de fertilizantes e automobilística se destacam pelo número elevado de inovações introduzidas, enquanto as indústrias têxtil e extrativa apresentam uma taxa bastante baixa de inovações, como mostrado pelos dados da Pesquisa de Inovação Tecnológica (PINTEC) de 2005, do Instituto Brasileiro de Geografia e Estatística (IBGE). Logo, uma questão relevante é entender os determinantes dessa heterogeneidade – seria ela explicada por características específicas (técnicas) dos setores, ou haveria também uma explicação econômica?

Os economistas conjecturam há algum tempo que a competição que a firma enfrenta no mercado determina suas estratégias em geral, e o gasto em P&D em particular. Certos mercados têm na diferenciação de produto uma estratégia importante de competição. Mercados de produtos *commodities* se caracterizam pela competição via preço, onde inovação de processo tem um papel importante. Além disso, diferentes intensidades de competição podem gerar incentivos diferentes à inovação. Schumpeter (1950) afirmava que o tamanho da firma, ou o seu poder de mercado, determinava positivamente o incentivo à inovação. Firms com elevado poder de mercado inovariam mais que outras com menos poder.

A compreensão da relação entre competição (ou seu inverso, poder de mercado) e inovação permite prever o impacto de medidas de incentivo, como a Lei do Bem, em diferentes indústrias. Isso possibilita uma adequação e melhora dessas

1. Segundo dados do Ministério da Ciência e Tecnologia (MCT) (<<http://www.mct.gov.br/index.php/content/view/7969.html>>), em torno de 2007 e 2008, o Brasil investiu 1,09% do Produto Interno Bruto (PIB) em P&D; a Argentina, 0,51%; e o México, 0,38%. O país com o maior investimento/PIB é o Japão, com 3,42%, seguido da Coreia do Sul, 3,37%; e dos Estados Unidos, 2,77%.

medidas. Por exemplo, incentivos fiscais à inovação em um setor pouco dinâmico e inovador podem ser inócuos se este for caracterizado por intensa competição. Nesse caso, talvez a melhor solução não seja uma política fiscal de incentivo à inovação, mas sim uma política de consolidação do setor.

Também existe um grande interesse em trabalhos empíricos sobre a relação entre competição e inovação na medida em que eles permitem “testar” os diferentes desdobramentos teóricos daquilo que Cohen, Levin e Mowery (1987) batizaram de “hipótese schumpeteriana”: inovação cresce com a concentração ou o poder de mercado. Uma vez que aqui nós utilizamos explicitamente o modelo desenvolvido por Aghion *et al.* (2002), é particularmente ilustrativo acompanharmos o percurso da hipótese schumpeteriana dentro da teoria do crescimento endógeno. Com efeito, nos chamados “modelos de primeira geração” (de crescimento endógeno), como os *quality ladders* de Grossman e Helpman (1991) ou Aghion e Howitt (1992), havia uma relação positiva e monótona entre poder de mercado e inovação. Já em Aghion *et al.* (2002), encontramos um novo resultado teórico: numa certa indústria, até um certo ponto, uma redução na concorrência tem impacto positivo sobre a taxa de inovação; mas aumentos sucessivos no poder de mercado acabarão reduzindo a taxa de inovação, caracterizando-se assim uma relação em U-invertido.² Desse modo, quando em nosso trabalho empírico olhamos para a indústria brasileira, é imediato nos perguntarmos se ela está mais próxima do mundo retratado nos modelos de 1ª geração, ou do mundo onde vale o U-invertido.

A maioria dos trabalhos aplicados à indústria brasileira, que revisamos na seção seguinte, utiliza a base de dados da PINTEC, compreendendo cerca de 12 mil observações (empresas) na sua edição de 2005, ao passo que a pesquisa IBMEC-Sensus 2008 é muito menor – apenas 371 observações. Assim, é importante justificar a escolha da nossa base: em primeiro lugar, a pesquisa IBMEC-Sensus tem como diferencial o fato de incluir também empresas agrícolas e do setor de serviços, e não só da indústria de transformação. Em segundo lugar, nossa base traz informações sobre um fundamento importante: a elasticidade da demanda dos diversos setores, ou bens “carro-chefe” das empresas. Essa segunda característica da nossa base está relacionada com um aspecto metodológico importante, que distingue nosso trabalho dos demais revisados na seção seguinte: a escolha da medida de poder de mercado.

Os trabalhos aplicados à indústria brasileira, e que investigam a relação entre inovação (ou gasto em P&D) e poder de mercado, utilizam sempre o índice de concentração de Hinfredahl-Hirschman (IHH) como medida de poder de mercado. Já nós partimos do modelo teórico de Aghion *et al.* (2002) para fornecer

2. Para maiores detalhes sobre esses modelos, remetemos o leitor à seção 3.

uma justificativa teórica para a especificação econométrica, e em consonância com esse modelo adotamos como medida de poder de mercado o índice de Lerner (IL). O IHH é uma medida pouco robusta, dependendo fortemente da definição do mercado de um certo produto, seus limites geográficos etc. Com efeito, pela teoria dos mercados contestáveis (BAUMOL, 1982), sabemos que um alto IHH pode coexistir com uma baixo IL, por exemplo, porque um monopolista atua num mercado onde as importações são livres, embora a quantidade importada seja zero.

Do ponto de vista puramente empírico, também encontramos uma justificativa para usar o IL, em vez do IHH. Uma vez que todas as empresas respondentes da pesquisa IBMEC-Sensus encontram-se listadas no *Balanço Anual da Gazeta Mercantil*, uma base muito mais ampla, com mais de 6 mil observações, nós usamos essa última base para computar o IHH e o IL para setores a 2 e a 3 dígitos de desagregação. Como seria de se esperar, encontramos uma correlação positiva entre essas duas medidas; mas essa correlação é relativamente baixa (35% a 2 dígitos e 22% a 3 dígitos), e por isso ao trabalharmos com o IL acreditamos estar trazendo alguma informação nova em relação aos trabalhos que já haviam utilizado o IHH para a indústria brasileira.

Assim, para Aghion *et al.* (2002), a variável explicativa adequada para se testar a hipótese schumpeteriana sobre a relação entre inovação (ou gasto em P&D) e poder de mercado é o IL. O IL refletiria inversamente o grau de *product market competition* (PMC), uma característica da “estrutura” de mercado. Contudo, como descrito por Schmalensee (1989), sabe-se em organização industrial que os trabalhos empíricos do paradigma “estrutura-conduta-desempenho” apresentam problemas; basicamente porque a “estrutura” (no caso, o índice de concentração ou o IL), variável explicativa desse paradigma, é na verdade endógena. Na literatura de crescimento endógeno, por exemplo, nos modelos de *quality ladder* de Aghion e Howitt (1992) ou de Grossman e Helpman (1991), isso fica ainda mais claro: firmas inovadoras (um aspecto de “desempenho”) são capazes de fazer *limit-pricing*, jogando os concorrentes para fora do mercado, e tornando-se monopolistas. A fim de resolver esse problema de endogeneidade, Aghion *et al.* (2002) utilizam como variáveis instrumentais uma série de leis e medidas de regulação da concorrência adotadas no Reino Unido nos anos 1970 e 1980, que teriam impactado exogenamente (isto é, sem serem elas mesmas afetadas pela *performance* inovadora) os ILs dos diversos setores. Neste trabalho usaremos uma estimativa da elasticidade-preço da demanda percebida pela firma, como um instrumento para o poder de mercado medido pelo IL. Essa informação é obtida da pesquisa IBMEC-Sensus, e constitui o principal diferencial dessa pesquisa para o propósito deste trabalho.

Quanto aos resultados obtidos, eles divergem de evidências anteriores para a indústria brasileira, e concordam parcialmente com os de Aghion *et al.* (2002): uma relação em U-invertido (*inverted U*) entre gasto em P&D e poder de mercado

é estatisticamente significativa quando usamos o IL, mas não quando usamos o IHH. Entretanto, quando usamos o inverso da elasticidade da demanda como variável instrumental, verificamos que, embora esse seja um instrumento forte para explicar o IL, a relação em U-invertido deixa de ser significativa. Comparando nossas estimativas de modelos OLS e *tobit* com nossas estimativas de modelos *probit*, também concluímos que o poder de mercado (IL) é significativo em explicar o montante gasto em P&D (mesmo controlando para o tamanho da firma e outras variáveis como oportunidade tecnológica, perfil exportador etc.), mas não a probabilidade de uma firma fazer P&D.³

O restante deste trabalho está organizado da seguinte maneira: a seção 2 revê a literatura empírica sobre a relação entre inovação (ou gasto em P&D) e poder de mercado, com ênfase na literatura aplicada ao caso brasileiro; a seção 3 traz uma breve e informal descrição do modelo de Aghion *et al.* (2002), a fim de dar ao leitor uma intuição sobre a relação em *inverted U*; a seção 4 apresenta as bases de dados que utilizamos: IBMEC-Sensus 2008 e *Balanço Anual da Gazeta Mercantil*, com ênfase nas distribuições setoriais do IL e do IHH; a seção 5 traz os resultados econométricos; a seção 6 conclui.

2 UMA BREVE REVISÃO DA LITERATURA

Na literatura de Organização Industrial (OI) empírica, Cohen, Levin e Mowery (1987) batizaram de “hipóteses schumpeterianas”: *i*) inovação cresce mais do que proporcionalmente com o tamanho das firmas; e *ii*) inovação cresce com a concentração ou o poder de mercado. Nos últimos anos, uma série de estudos procurou testar essas hipóteses para a indústria brasileira, usando diferentes metodologias, diferentes bases de dados e diferentes variáveis como *proxies* para inovação, tamanho, poder de mercado etc.

Jensen, Menezes-Filho e Sbragia (2004) testaram a hipótese (*i*) usando um painel não balanceado de 420 empresas durante cinco anos (1994 a 1998), com dados da Associação Nacional de Pesquisa e Desenvolvimento das Empresas Inovadoras (Anpei). Eles obtiveram uma elasticidade do gasto em P&D em relação ao faturamento igual a 0.65, portanto rejeitando a hipótese de que “inovação” cresce mais que proporcionalmente ao tamanho das firmas.

Kannebley Jr., Porto e Pazzello (2005) dispõem de uma base bem mais ampla, a da PINTEC 2000, com 10 mil firmas. O método não paramétrico utilizado,

3. Este último resultado guarda uma analogia interessante com os testes da hipótese schumpeteriana feitos por Cohen, Levin e Mowery (1987, p. 1), que concluem que “*Only one size-related effect withstands scrutiny: although neither measure of size influences the behavior of R&D performers [isto é, o montante gasto em P&D], business unit size does affect the probability of conducting R&D*”. No nosso caso, usando a amostra IBMEC-Sensus 2008, concluímos exatamente o contrário sobre o “poder de mercado” (IL), controlando, é claro, para o tamanho das empresas.

classification trees, permite determinar qual é a combinação (ordenada) de variáveis independentes que melhor prevê uma variável dependente binária, que assume valor 1 quando, de acordo com o questionário da PINTEC 2000, uma firma introduziu alguma inovação no período de análise, e valor zero, caso contrário. Usando a definição mais ampla de inovação (incluindo também a possibilidade de a firma ter feito algo que representa uma inovação apenas para ela mesma), descobre-se que tamanho é a segunda variável explicativa mais importante, depois de perfil exportador. Usando a definição estrita de inovação (inovação em relação ao mercado brasileiro), descobre-se ou que o tamanho da firma está longe de ser a variável mais importante para explicar o comportamento inovador, ou que não há uma relação monótona entre tamanho e inovação.

Pires-Alves e Rocha (2008) questionam esse último resultado de Kannebley Jr., Porto e Pazello (2005), apontando para o fato de que muitas firmas pequenas da base PINTEC 2000 responderam terem feito alguma inovação ao mesmo tempo em que tiveram gasto zero em P&D. Adotando, assim, uma variável dependente binária que vale 1 quando a firma teve gasto positivo em P&D e zero, caso contrário, Pires-Alves e Rocha (2008) rodam com a PINTEC diversos modelos *probit* que sugerem os seguintes fatos estilizados: *iii*) firmas grandes têm maior probabilidade de fazerem P&D; e *iv*) a probabilidade de uma firma fazer P&D cresce com o poder de mercado, que é medido através do índice de concentração IHH. Contrariando uma tendência já apontada por Cohen, Levin e Mowery (1987), esses resultados com variáveis estruturais resistem à introdução de *dummies* de setor e de variáveis de oportunidade tecnológica. Pires-Alves e Rocha (2008) também realizam um teste mais direto das hipóteses *i*) e *ii*) citadas anteriormente, tomando como variável dependente em modelos *tobit* a intensidade de P&D (gasto em P&D/receita de vendas). Basicamente, para a indústria brasileira como um todo, os resultados apontam para intensidade de P&D dependendo positivamente de tamanho da firma (medido pelo número de funcionários) e poder de mercado.

Na literatura estrangeira é importante mencionar Scherer (1967), a quem talvez se possa atribuir a “descoberta” da relação em U-invertido: utilizando dados da indústria manufatureira dos Estados Unidos em 1960, e controlando para características de produto, oportunidade tecnológica e setores, Scherer descobriu que a proporção de funcionários empregados em P&D era uma função côncava do grau de concentração da indústria.

Uma boa resenha de outros trabalhos que obtiveram ora uma relação positiva, ora uma relação negativa, entre P&D e poder de mercado pode ser encontrada em Tingvall e Poldahl (2006), que se propõem eles mesmos a testar o modelo de Aghion *et al.* (2002) utilizando dados da indústria manufatureira sueca de 1990 a

2000: elegendo o gasto em P&D como a medida mais adequada para inovação,⁴ esses autores encontram uma relação em U-invertido quando o poder de mercado numa certa indústria é medido pelo IHH, mas não quando ele é medido pelo *mark-up* de preço das firmas individuais (uma medida que corresponderia a um IL para cada firma, e não ponderado por setor, como nós fazemos neste trabalho).

Convém lembrar que, na parte empírica de seu trabalho, Aghion *et al.* (2002) utilizam um painel (de 1968 a 1996) de firmas do Reino Unido, tendo como variável dependente (para inovação) o número de patentes registradas no United States Patent and Trademark Office (USPTO), e como principal variável explicativa (para poder de mercado) um IL ponderado por setor idêntico ao que empregamos aqui e que é descrito em detalhes na seção 4.

Agora, é importante situar nossa contribuição em relação a essa literatura. Como fica claro na introdução deste trabalho, nosso foco está na relação entre gasto em P&D (tanto em nível quanto em probabilidade) e poder de mercado, de modo que Pires-Alves e Rocha (2008), Tingvall e Poldahl (2006) e obviamente Aghion *et al.* (2002) são nossos interlocutores mais importantes. No caso da indústria brasileira, é interessante ressaltar que enquanto Pires-Alves e Rocha (2008) obtêm uma relação em U-invertido entre gasto em P&D e o IHH usando a base da PINTEC 2005, na seção 5 veremos que na base IBMEC-Sensus essa relação não é significativa usando o IHH, e sim usando o IL. Além disso, comparado aos outros trabalhos sobre a indústria brasileira aqui citados, nosso trabalho traz (na subseção 5.1) uma discussão metodológica e um teste mais cuidadosos da relação em U-invertido, baseados em Lind e Mehlun (2010).⁵ Outro ponto que merece destaque em nosso trabalho é o tratamento do problema da (possível) endogeneidade entre gasto em P&D e poder de mercado, que está ausente em Pires-Alves e Rocha (2008): mostramos aqui que uma estimativa da elasticidade da demanda ao nível do setor (3 dígitos de desagregação), e obtida a partir de uma resposta estimulada do questionário da pesquisa IBMEC-Sensus 2008, é um instrumento forte para explicar o IL. No entanto, os resultados de segundo estágio apontam para uma simultaneidade entre gasto em P&D e poder de mercado.

Outras questões mencionadas nesta revisão da literatura são abordadas apenas marginalmente, quando olhamos para os controles usados em nossas regressões: tamanho da firma, setor da firma, se a firma é exportadora etc.⁶

4. Segundo os autores, medidas como número de patentes seriam inadequadas devido às grandes diferenças de propensão a patentear entre os setores.

5. Agradecemos ao editor da revista PPE por essa sugestão.

6. Apenas a título de curiosidade, quando rodamos o mesmo tipo de regressão que Jensen, Menezes-Filho e Sbragia (2004), isto é, $\ln(\text{gasto em P\&D}_i) = \beta_1 \cdot \ln(\text{receita}_i) + \beta_2 \cdot \text{dummy de setor}_i + \varepsilon_i$, onde i denota a firma, encontramos uma elasticidade do gasto em P&D em relação ao faturamento igual a 0.42 na amostra IBMEC-Sensus 2008, portanto também rejeitando a hipótese schumpeteriana de que "inovação" cresce mais que proporcionalmente ao tamanho das firmas.

3 INOVAÇÃO E CONCORRÊNCIA: UMA RELAÇÃO EM U-INVERTIDO

Existem evidências empíricas de que gasto em P&D e poder de mercado têm uma relação em U-invertido: Pires-Alves e Rocha (2008), por exemplo, usam em suas regressões o IHH e o IHH elevado ao quadrado como variáveis explicativas, obtendo coeficientes estimados que são respectivamente positivo e negativo. O objetivo desta seção é dar a essa regularidade empírica uma base teórica, base que mais tarde, na seção 5, iremos usar para justificar as especificações de nossos modelos econométricos. Para isso, usamos o arcabouço desenvolvido originalmente por Aghion *et al.* (2002).

Em modelos schumpeterianos tradicionais de crescimento endógeno, as firmas inovam (lançando novos produtos ou novos métodos de produção) esperando com isso capturar renda de monopólio. Líderes tecnológicos jogam firmas mais atrasadas para fora do mercado via *limit pricing*. Esse comportamento dos líderes só é possível quando os bens são substitutos perfeitos. Por outro lado, devido a vazamentos de conhecimento (*knowledge spillovers*) que afetam o cálculo de ganhos incrementais *versus* custos incrementais, líderes tecnológicos não têm incentivo para inovar ainda mais, acima do seu próprio “estado da arte”. Nesse contexto, um aumento da concorrência (PMC) provoca uma queda na taxa de inovação da indústria. Era isso o que acontecia nos modelos *quality ladder* de 1ª geração, como os de Grossman e Helpman (1991) ou Aghion e Howitt (1992). Comparado com isso, a grande mudança de hipótese em Aghion *et al.* (2002) é que dentro da mesma indústria “variedades” do mesmo bem são substitutas imperfeitas, de modo que retardatários tecnológicos têm lucro maior do que zero, o que por sua vez fará com que mudanças na competição afetem de forma mais complexa os incentivos à inovação.

Nesta versão de nosso trabalho, não apresentaremos formalmente o arcabouço de Aghion *et al.* (2002).⁷ Passaremos apenas a intuição por trás do novo resultado teórico: numa certa indústria, até um certo ponto, uma redução na concorrência tem impacto positivo sobre o crescimento via inovação; mas aumentos sucessivos no poder de mercado acabarão reduzindo a taxa de inovação, caracterizando-se assim a relação em U-invertido.

Primeiramente, note que nessa nova situação (quando, dentro da mesma indústria, “variedades” do mesmo bem são substitutas imperfeitas, de modo que retardatários tecnológicos têm lucro maior que zero) uma certa indústria pode se encontrar em dois diferentes estados: uma indústria nivelada (*leveled* ou *neck and neck*) é uma na qual todas as firmas têm a mesma tecnologia. Chamemos os lucros correspondentes de π_0 . Ou então uma indústria pode ser tal que há um

7. O leitor interessado pode ir diretamente à fonte original, ou pedir-nos a versão completa do nosso *working paper*.

líder tecnológico que está um passo além dos seguidores, isto é, uma indústria desnivelada (*unleveled*). Chamemos os lucros correspondentes de π_1 (do líder) e π_{-1} (dos seguidores). Faz sentido assumirmos $\pi_{-1} < \pi_0 < \pi_1$.

A seguir, chamemos de N a intensidade agregada de inovação.⁸

$$N = \underbrace{\mu_L}_{\substack{\text{proporção} \\ \text{de indústrias} \\ \text{leveled}}} \cdot n_0 + \underbrace{\mu_U}_{\substack{\text{proporção} \\ \text{de indústrias} \\ \text{unleveled}}} \cdot n_{-1} \Rightarrow dN \cong \mu_L \cdot dn_0 + \mu_U \cdot dn_{-1} \quad (1)$$

onde n_0 e n_{-1} denotam gastos em P&D, ou investimento em inovação, por firmas niveladas e seguidoras, respectivamente. Recorde que, nesta classe de modelos, líderes não têm incentivo para inovar em cima do seu próprio nível tecnológico.

Agora, como será que os pesos μ_L e μ_U dependem, por sua vez, da PMC? Quando a PMC é baixa, o diferencial ($\pi_1 - \pi_0$) é pequeno, e portanto o esforço n_0 para se desnivelar e virar líder é baixo, de modo que uma indústria nivelada leva um longo tempo para se tornar desnivelada. Por outro lado, sob uma baixa PMC o diferencial ($\pi_0 - \pi_{-1}$) é grande, e portanto um seguidor numa indústria desnivelada tem um grande incentivo para inovar (alto n_{-1}), de modo que uma indústria desnivelada tende rapidamente a passar a nivelada. Assim, sob uma baixa PMC, uma indústria típica vai passar a maioria do tempo sendo uma indústria nivelada, o que significa um grande peso μ_L e um pequeno μ_U . De modo que nesse caso o que predomina é o efeito de mudanças na PMC sobre n_0 : um aumento na PMC vai reduzir π_0 e elevar o diferencial ($\pi_1 - \pi_0$), despertando o incentivo para uma firma nivelada inovar, isto é, aumentando n_0 . É isso que Aghion *et al.* (2002) chamam de *escape-competition effect*. Resumindo, quando a PMC é baixa, mais PMC leva a mais inovação no agregado.

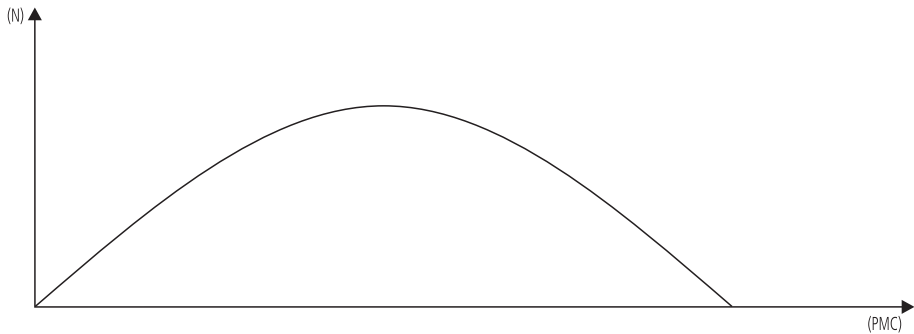
Pelo contrário, quando a PMC é alta, o diferencial de lucros ($\pi_1 - \pi_0$) também é alto e portanto alguém que está numa indústria nivelada tem um alto incentivo para inovar \rightarrow alto n_0 , e uma indústria nivelada leva pouco tempo para passar a desnivelada. Por outro lado, quando a PMC é alta, é baixo o diferencial ($\pi_0 - \pi_{-1}$), e portanto um seguidor numa indústria desnivelada não tem grande incentivo para inovar \rightarrow baixo n_{-1} , e uma indústria desnivelada leva um longo tempo para passar a nivelada. Assim, sob uma alta PMC, uma indústria típica vai passar a maioria do tempo sendo uma indústria desnivelada, o que significa um grande peso μ_U e um pequeno μ_L . De modo que nesse caso o que predomina é o efeito de mudanças

8. No modelo original, essa "intensidade agregada de inovação" é o montante de recursos reais aplicados em P&D, tendo, assim, uma correspondência perfeita com a nossa variável dependente gasto em P&D.

na PMC sobre n_{-1} : um aumento na PMC vai reduzir π_0 e com ele o diferencial $(\pi_0 - \pi_{-1})$, reduzindo o incentivo para um seguidor inovar, isto é, reduzindo n_{-1} . É isso que Aghion *et al.* (2002) chamam de *Schumpeterian effect*. Resumindo, quando a PMC já é alta, mais PMC leva a menos inovação no agregado.

GRÁFICO 1

A relação em U-invertido entre PMC e inovação agregada (N)



Fonte: Elaboração dos autores.

Uma questão importante para a análise empírica é: mas como a PMC é parametrizada no modelo? Aghion *et al.* (2002) assumem que em cada indústria j existem duas variedades do mesmo bem, A e B , produzidas por dois oligopolistas, e que o argumento que entra na função utilidade dos consumidores é x_j , uma espécie de agregado dessas duas variedades, dado por:

$$x_j = \left(x_{Aj}^{\alpha_j} + x_{Bj}^{\alpha_j} \right)^{\frac{1}{\alpha_j}} \quad (2)$$

onde $0 < \alpha_j < 1$, e x_A e x_B são as quantidades produzidas pelos oligopolistas.

Na expressão (2), α_j é o parâmetro que captura a *elasticidade de substituição* entre as duas variedades: quando α_j é igual a 1, as duas variedades são substitutos perfeitos, e a PMC é máxima. Quando α_j é bem menor do que 1, é grande o poder de monopólio na indústria j . Assim, um aumento na PMC é parametrizado através de um aumento em α_j . A seguir, Aghion *et al.* (2002) demonstram que η_{ij} , a *elasticidade-preço da demanda* com que se defronta o oligopolista i na indústria j , é dada por:

$$\eta_{ij} = \frac{1 - \lambda_{ij} \cdot \alpha_j}{1 - \alpha_j} \quad (3)$$

onde $0 < \lambda_{ij} < 1$ é o *market-share* do oligopolista i na indústria j . Com isso, é imediato ver que $\partial\eta/\partial\alpha > 0$, de modo que, com *market-shares* constantes, um aumento na PMC (isto é, em α) leva a um aumento na elasticidade da demanda. Uma vez que o IL é dado pelo inverso da elasticidade-preço da demanda, também é imediato que:

$$IL_{ij} = \frac{p_i - c_i}{p_i} = \frac{1 - \alpha_j}{1 - \alpha_j \cdot \lambda_{ij}} \Rightarrow \frac{\partial(\text{Lerner Index})}{\partial\alpha_j} < 0 \quad (4)$$

onde c_i é o custo unitário do oligopolista i . Assim, dados os *market-shares* da indústria j , podemos concluir que, quanto maior a PMC (α_j), menores serão os ILs nessa indústria. Embora não possamos falar numa relação 1 para 1 entre o IL e a PMC, e nem mesmo que essa relação é independente da medida de concentração da indústria (λ s), Aghion *et al.* (2002, p. 13) concluem que: “*The following considerations, suggest that the average Lerner index of a random sample of firms in an industry should be a decreasing function of the measure of PMC, even after taking into account the effect of a on R&D intensities and therefore on average market shares*” (grifo nosso).

Por isso, seguindo Aghion *et al.* (2002), na nossa análise econométrica vamos usar uma média amostral ponderada dos ILs de cada indústria como uma medida da PMC naquela indústria.⁹

4 SOBRE OS DADOS

Apesar de ter uma base pequena (371 empresas), o questionário da pesquisa IBMEC-Sensus 2008 é bastante amplo e compreensivo, abrangendo desde aspectos financeiros e organizacionais das empresas, até questões relativas a estrutura e poder de mercado e, é claro, tecnologia e P&D.

Para a nossa análise econométrica, as seguintes questões da pesquisa forneceram dados:¹⁰ gastos totais com P&D em 2006 (em R\$ correntes); sensibilidade da quantidade vendida a aumentos no preço; se o produto da empresa está mais

9. Observando a tabela 1, constatamos que mesmo a 2 dígitos de desagregação há uma grande variedade de IHHs (e portanto de distribuições dos *market-shares*) entre os diversos setores. Junto com a expressão (4), isso significa que nem toda a variedade de ILs entre setores consiste em diferenças de “poder de mercado” (ou da sua medida inversa, a PMC), quando essa última é interpretada estritamente como refletindo os microfundamentos de preferências do modelo (parâmetro α). Já *dentro de um dado setor*, o modelo de Aghion *et al.* (2002) implica que pode sim haver heterogeneidade de *market-shares*, mas não de poder de mercado, uma vez que há um único α , para cada setor (bem) j .

10. Para maiores detalhes, ver o apêndice A, com uma lista comentada das questões da pesquisa IBMEC-Sensus que utilizamos.

para uma *commodity* ou um bem diferenciado; porcentagem da receita advinda de exportações.¹¹ As variáveis de regressão correspondentes foram chamadas respectivamente de: *gastopd*; *invelast* (inverso da elasticidade); *commodity* (*dummy*); *reexports*.

Além dessas questões da pesquisa IBMEC-Sensus 2008, utilizamos a base do *Balanço Anual* da *Gazeta Mercantil* (2006),¹² para obter dados sobre receita líquida e lucro operacional das empresas. Note que todas as empresas da base IBMEC-Sensus 2008 encontram-se listadas no *Balanço Anual* da *Gazeta*, de modo que a base IBMEC-Sensus é na verdade um subconjunto da base da *Gazeta*. Numa primeira etapa, de cada uma das 180 empresas que declararam gastos em P&D na pesquisa IBMEC-Sensus, extraímos o dado de *receita líquida*, que usaremos como *controle para “tamanho”* em nossas regressões (nome da variável de regressão: *rec liq*).

A etapa seguinte foi computar os ILs e IHHs para cada setor a 2 e a 3 dígitos de desagregação,¹³ a partir dos dados de receita líquida e lucro operacional. Para cada firma em nossa amostra IBMEC-Sensus identificamos o setor a 2 dígitos e o subsetor a 3 dígitos SIC (Standard Industrial Classification, utilizada pelo *Balanço* da *Gazeta Mercantil*) a que esta pertence. Seleccionamos então todas as firmas na base da *Gazeta Mercantil* que pertencem a este setor (ou subsetor), como base de cálculo para os índices. É importante notar que, dessa maneira, a cada observação da variável dependente gasto em P&D, proveniente da pequena base IBMEC-Sensus 2008, associaremos em nossas regressões uma medida do IL e uma do IHH do setor correspondente, só que provenientes de uma base muito maior (*Gazeta*, com mais de 6 mil observações). Algo idêntico é feito por Aghion *et al.* (2002, p. 31): “*We are able to measure the Lerner index using information on all UK firms in Datastream, not only the sample for which we also observe patents. (...) In computing this index, we use the entire sample of Stock Market Listed firms in each industry, not only those in the patenting subsample*”.

Nossos IHHs e ILs amostrais são computados da seguinte maneira:

$$IHH_j = \sum_{i \in j} (\text{market share}_i)^2, \text{ para a indústria/setor } j$$

11. Recorde, da seção 2, que na *classification tree* de Kannebley Jr., Porto e Pazello (2005) uma variável *dummy* do tipo exportava = 1, não exportava = 0 aparece como um dos principais determinantes do comportamento inovador das empresas.

12. A escolha do ano 2006 deve-se a que o questionário da pesquisa IBMEC-Sensus 2008 correu durante 2007 entre as empresas, de modo que as respostas tinham como base o ano 2006.

13. No *Balanço Anual* da *Gazeta Mercantil*, o nível de desagregação a 1 dígito é chamado de NOME DA ATIVIDADE; a 2 dígitos é chamado de SETOR; a 3 dígitos é chamado de SUBSETOR.

onde:

$$market\ share_i = \frac{receita\ líquida_i}{\sum_{k \in j} receita\ líquida_k} \quad (5)$$

Essa é a forma usual de se calcular o IHH, portanto, dispensando mais comentários. Já quanto ao IL, usamos a seguinte *proxy*:

$$IL_j = \frac{\sum_{i \in j} lucro\ operacional_i}{\sum_{i \in j} receita\ líquida_i}, \text{ para a indústria/setor } j. \quad (6)$$

Ao utilizarmos esse método de construção do IL, estamos calculando um índice médio ponderado para cada setor.¹⁴ Também como Aghion *et al.* (2002), usamos o lucro operacional para computar o IL.¹⁵

A seguir, vejamos as distribuições setoriais do IL e do IHH (tabela 1):

TABELA 1
ILLs e IHHs para setores a 2 dígitos – 2006

Lista de setores a 2 dígitos	IL	IHH
Agricultura, pecuária e cooperativas	0.03	0.03
Alimentos, bebidas e fumo	0.05	0.06
Atacadistas	0.02	0.31
Comércio exterior	0.03	0.05
Comércio varejista	0.02	0.09
Construção e engenharia	0.07	0.12
Distribuidores de veículos e autopeças	0.03	0.02
Elerodomésticos	0.02	0.10
Energia elétrica	0.11	0.06
Equipamentos elétricos	0.08	0.08
Farmacêuticos	0.11	0.08
Higiene, limpeza e cosméticos	0.14	0.29

(continua)

14. Médio ponderado se refere ao fato de que firmas maiores de um setor têm um maior peso no índice.

15. Uma alternativa seria usar, em vez do lucro operacional, o EBITDA – sigla em inglês para *earnings before interest, taxes, depreciation and amortization* – mas, no caso brasileiro, acreditamos que essa última medida incluiria uma variabilidade indesejada de custos financeiros entre empresas. No caso do Reino Unido, onde o acesso ao crédito é muito mais uniforme, as duas medidas seriam equivalentes.

(continuação)

Lista de setores a 2 dígitos	IL	IHH
Madeira e móveis	0.26	0.09
Mecânica	0.04	0.05
Metalurgia	0.11	0.03
Mineração	0.31	0.34
Minerais não metálicos	0.10	0.04
Papel e celulose	0.09	0.08
Petróleo e gás	0.17	0.39
Plásticos e borracha	0.06	0.06
Química	0.06	0.05
Saneamento e limpeza	0.09	0.08
Saúde	0.03	0.02
Serviços especializados, esportes, turismo e lazer	0.10	0.04
Siderurgia	0.25	0.09
Tecnologia da informação	0.11	0.71
Telecomunicações	0.05	0.08
Têxtil, couro, calçados e confecções	0.02	0.02
Transporte e logística	0.03	0.03
Veículos e autopeças	0.04	0.13

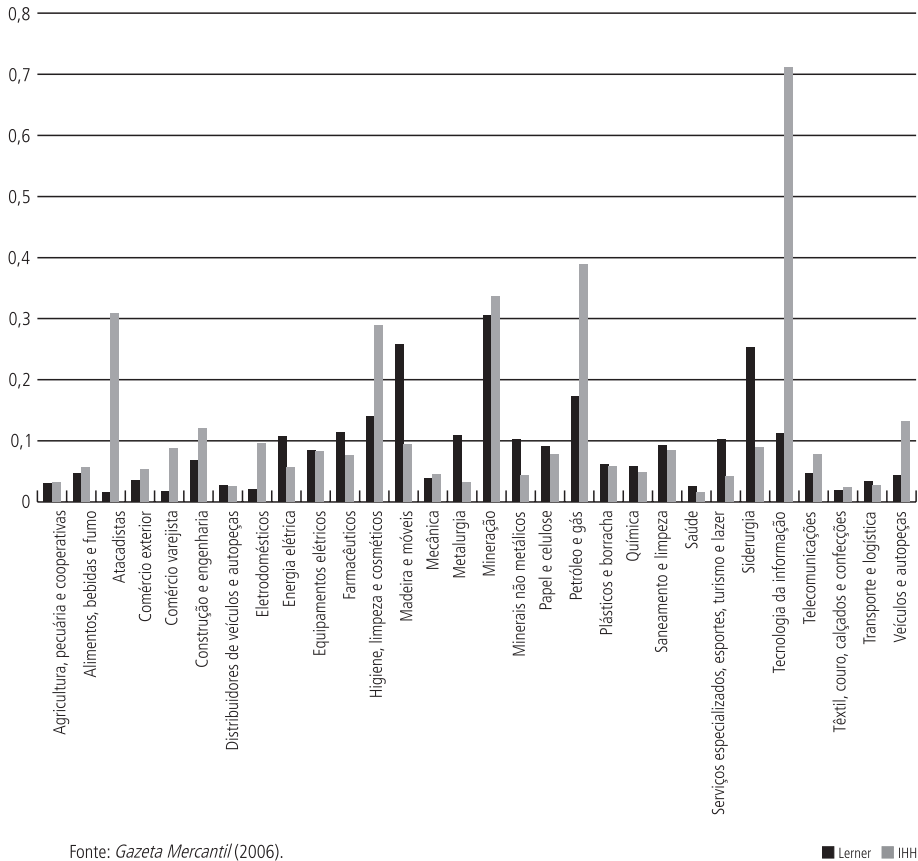
Fonte: *Gazeta Mercantil* (2006).

Para 2 dígitos de desagregação, o coeficiente de correlação entre o IL e o IHH é 0.3460, o que indica haver uma previsível correlação positiva, mas consideravelmente imperfeita.

No apêndice B, trazemos uma tabela com os IHHs e ILs a 3 dígitos. Nesse nível de desagregação, a correlação é ainda menor, 0.2196. Concluímos assim que, do ponto de vista empírico, ao trabalharmos com o IL estamos capturando informações novas, não contidas no IHH, já usado por outros trabalhos aplicados ao caso brasileiro. Além disso, como vimos na seção 1, Aghion *et al.* (2002) argumentam que o IHH é uma medida pouco robusta, dependendo da definição do que é o mercado de um certo produto, seus limites geográficos etc.¹⁶ Assim, utilizamos o IL como uma medida do poder de mercado da firma, ou do inverso da PMC.

16. A esse respeito, ver Aghion *et al.* (2002, nota 21, p. 29), com um interessante exemplo para a indústria farmacêutica no Reino Unido, cujas empresas domésticas possuem grandes *market-shares* no próprio mercado doméstico, mas pequenos no mercado mundial.

GRÁFICO 2
IHHs e ILs para setores a 2 dígitos – 2006



Fonte: *Gazeta Mercantil* (2006).

Uma questão relevante foi a escolha do nível de desagregação mais adequado para a análise de regressão. Em princípio, quanto maior o nível de desagregação, mais fielmente o IL estará refletindo a PMC num mercado específico de produto. Por isso, optamos por trabalhar com o IL e o IHH a 3 dígitos de desagregação (nome das variáveis de regressão: *lerner* e *IHH*).

A última variável que extraímos do *Balanço Anual* da *Gazeta Mercantil* é uma *dummy* de setor a 1 dígito de desagregação, que usaremos como controle, motivados pelo fato estilizado apontado por Cohen, Levin e Mowery (1987): resultados com variáveis estruturais (no nosso caso, principalmente o IL) não resistiriam à introdução de *dummies* de setor e de variáveis de oportunidade tecnológica (nome das variáveis de regressão: *dagro*, *dserv*, *dcomerc* (usamos indústria como base).

Terminamos esta seção com algumas *estatísticas descritivas* sobre as nossas variáveis de regressão (tabela 2):

TABELA 2
Estatísticas descritivas

	Observação	Média	Desvio-padrão	Mínimo	Máximo
<i>gastopd</i>	180	1197547	7142769	0	80000000
<i>lerner</i>	180	0.075	0.094	0	0.5
<i>lerner</i> ²	180	0.015	0.039	0	0.25
<i>IHH</i>	180	0.175	0.2042	0.0247	0.9746
<i>IHH</i> ²	180	0.0721	0.1695	0.00061	0.9499
<i>inelast</i>	154	0.66	0.376	0.059	1
<i>inelast</i> ²	154	0.576	0.459	0.003	1
<i>dagro</i>	180	0.072	0.26	0	1
<i>dcomerc</i>	180	0.183	0.388	0	1
<i>dserv</i>	180	0.544	0.499	0	1
<i>rec_liq</i>	180	94028000	516667282.5	37990	6789508000
<i>commodity</i>	155	0.4451	0.4985	0	1
<i>reexports</i>	135	8.27	23.2	0	100

Fonte: Cálculos dos próprios autores.

5 A RELAÇÃO EM U-INVERTIDO NA BASE IBMEC-SENSUS: RESULTADOS ECONÔMICOS

Vejamos agora os resultados das diferentes estimações realizadas para mostrar a relação entre gasto em P&D e poder de mercado. Como a pesquisa IBMEC-Sensus foi realizada uma única vez, durante o ano de 2007, e colhendo informações do ano 2006, todas as nossas regressões são *cross-sections*. Temos três modelos – Ordinary Least Squares (OLS), *tobit* e *probit* –, sendo que para os dois primeiros modelos usamos três especificações diferentes, conforme os controles incluídos. Nas tabelas 3 e 4, as estimações enumeradas de I a III são regressões lineares OLS, enquanto as colunas IV a VI mostram os resultados de um modelo *tobit* e a coluna VII de um *probit*. Em todas as estimações a variável dependente é o gasto em P&D (*gastopd*), com exceção do modelo *probit*, onde a dependente tem valor 1 se o gasto em P&D é maior que zero; e zero, caso contrário.

As variáveis explicativas são¹⁷ receita líquida (*rec_liq*), poder de mercado expresso pelo IHH (tabela 3), poder de mercado expresso pelo IL (*lerner*, tabela 4),

17. Para uma discussão mais detalhada das variáveis, como elas foram extraídas etc., ver seção 4.

dummies de setor a 1 dígito (*dagro* – agroindustrial, *dcom* – comércio e *dserv* – serviços, sendo a indústria a categoria base), a *dummy* chamada *commodity*, que assume valor 1 se a firma declarou que seu principal produto é uma *commodity* e zero, caso contrário; e *recexports*, que é a porcentagem das receitas da empresa obtida através de exportações.

TABELA 3
Resultados das regressões com poder de mercado \equiv IHH
(Variável dependente: *gastopd*)

	OLS			Tobit			Probit
	I	II	III	IV	V	VI	VII
<i>IHH</i>	8.68 <i>0.63</i>	0.23 <i>0.02</i>	0.15 <i>0.02</i>	22.41 <i>0.77</i>	9.15 <i>0.36</i>	12.80 <i>0.76</i>	1.95 <i>0.90</i>
<i>IHH</i> ²	-9.94 <i>-0.61</i>	-2.74 <i>-0.18</i>	-1.43 <i>-0.16</i>	-17.72 <i>-0.51</i>	-12.08 <i>-0.37</i>	-14.23 <i>-0.70</i>	-1.06 <i>-0.42</i>
<i>dagro</i>	-0.23 <i>-0.07</i>			-2.45 <i>-0.36</i>			-0.33 <i>-0.62</i>
<i>dcomerc</i>	-0.53 <i>-0.19</i>			-9.95 <i>-1.610*</i>			-1.12 <i>-2.464**</i>
<i>dserv</i>	2.35 <i>1.03</i>			-0.92 <i>-0.20</i>			-0.49 <i>-1.40</i>
<i>rec_liq</i>	0.00 <i>0.71</i>	0.00 <i>0.87</i>	5.74 <i>6.050***</i>	0.01 <i>0.98</i>	0.01 <i>1.15</i>	7.30 <i>4.151***</i>	0.00 <i>0.86</i>
<i>commodity</i>	1.80 <i>1.10</i>	2.16 <i>1.35</i>		2.32 <i>0.66</i>	3.26 <i>0.34</i>		-0.08 <i>-0.31</i>
<i>rec_export</i>	0.01 <i>0.19</i>	0.00 <i>-0.07</i>		0.05 <i>0.71</i>	0.09 <i>1.29</i>		0.01 <i>1.26</i>
<i>_cons</i>	-1.51 <i>-0.59</i>	0.42 <i>0.26</i>	0.74 <i>0.85</i>	-11.06 <i>-2.086**</i>	-12.19 <i>-3.122***</i>	-9.24 <i>-4.093***</i>	-0.24 <i>-0.61</i>
obs	112	112	180	112	112	180	112

Fonte: Elaboração dos autores.

Notas: Os números em itálico são estatísticas-t e z.

*significativo a 10%; **significativo a 5%; ***significativo a 1%.

De modo geral, a fim de testar a hipótese do U-invertido, rodamos regressões do tipo:

$$gastopd_i = \beta_0 + \beta_1 \cdot \text{poder de mercado}_i + \beta_2 \cdot (\text{poder de mercado}_i)^2 + \tilde{\beta}_3 \cdot \text{controles}_i + \varepsilon_i \quad (7)$$

esperando obter $\beta_1 > 0$ e $\beta_2 < 0$.

TABELA 4

Resultados: poder de mercado \equiv IL(Variável dependente: *gastopd*)

	OLS			Tobit			Probit
	I	II	III	IV	V	VI	VII
<i>lerner</i>	68.20 <i>2.74***</i>	60.10 <i>2.67***</i>	28.90 <i>2.14***</i>	108.00 <i>2.17***</i>	113.00 <i>2.51***</i>	62.20 <i>2.16***</i>	1.18 <i>0.28</i>
<i>lerner</i> ²	-123.00 <i>-2.28***</i>	-109.00 <i>-2.16***</i>	-54.20 <i>-1.66**</i>	-177.00 <i>-1.73**</i>	-186.00 <i>-1.95***</i>	-94.40 <i>-1.43*</i>	2.00 <i>0.19</i>
<i>dagro</i>	0.40 <i>0.12</i>			-0.73 <i>-0.11</i>			-0.24 <i>-0.45</i>
<i>dcomerc</i>	2.84 <i>1.08</i>			-2.16 <i>-0.37</i>			-0.72 <i>-1.65</i>
<i>dserv</i>	2.50 <i>1.21</i>			-1.26 <i>-0.30</i>			-0.60 <i>-1.82</i>
<i>rec_liq</i>	0.92 <i>0.17</i>	1.37 <i>0.25</i>	5.33 <i>5.59***</i>	5.96 <i>0.60</i>	5.66 <i>0.57</i>	6.25 <i>3.59***</i>	0.00 <i>0.88</i>
<i>commodity</i>	-1.59 <i>-1.01</i>	-1.67 <i>-1.08</i>		-2.02 <i>-0.61</i>	-2.01 <i>-0.61</i>		0.05 <i>0.21</i>
<i>rec_export</i>	0.00 <i>-0.12</i>	-0.02 <i>-0.64</i>		0.03 <i>0.39</i>	0.04 <i>0.54</i>		0.01 <i>1.19</i>
<i>_cons</i>	-0.80 <i>-0.23</i>	1.62 <i>0.58</i>	-0.70 <i>-0.91</i>	-9.54 <i>-1.36</i>	-10.80 <i>-1.77</i>	-11.20 <i>-5.51</i>	-0.21 <i>-0.38</i>
obs	112	112	180	112	112	180	112

Fonte: Elaboração dos autores.

Notas: Os números em itálico são estatísticas-t Student.

*significativo a 15%; **significativo a 10%; ***significativo a 5%.

Os resultados das nossas estimações quando o poder de mercado é expresso pelo IHH foram vistos na tabela 3.¹⁸

Embora os resultados sugeriram uma relação em U-invertido (uma vez que em todas as especificações e em todos os modelos o coeficiente associado ao IHH é positivo e o coeficiente associado a IHH² é negativo), em nenhum caso a relação se sustenta a níveis razoáveis de significância.

Com isso, partiu-se para a tabela 4, que apresenta os resultados das diferentes estimações realizadas para mostrar a relação entre gasto em P&D e poder de mercado, sendo esse representado pelo IL.

18. A fim de facilitar a interpretação dos resultados das regressões, nas tabelas 3, 4 e 5, as variáveis *gastopd* e *rec_liq* aparecem já divididas por 1.000.000.

Podemos resumir os resultados da tabela 4 nos seguintes pontos:

- 1) O coeficiente associado ao IL é positivo e o coeficiente associado a IL^2 é negativo em todos os modelos estimados via OLS e *tobit*, o que é compatível com uma relação na forma de U-invertido.
- 2) As regressões estimadas via OLS são todas significativas pelo menos a 10% de significância, com algumas especificações mostrando significância de 5%. Os resultados das estimações via *tobit* também se mostram significativos, alcançando significância de 5%. Por falta de significância, o modelo *probit* mostrou-se inadequado para explicar os dados e não será considerado na análise que se segue.
- 3) Os coeficientes do IL se mostraram bastante robustos e não mudam de forma significativa de acordo com a especificação da equação ou com o método de estimação (exceto no *probit*).

Tal como em Pires-Alves e Rocha (2008), tais resultados são robustos à introdução de *dummies* de setor. De modo geral, as demais variáveis de controle não se mostraram relevantes para obter o resultado compatível com a relação na forma de U-invertido. Mesmo assim, consideramos crucial controlar para o tamanho da firma através da inclusão da receita líquida nas regressões.

5.1 Testando a relação em U-invertido

Note, contudo, que até agora não podemos ainda concluir que a relação entre gasto em P&D e poder de mercado (medido pelo IL) tem realmente a forma de U-invertido. Afinal, o resultado obtido nas regressões da tabela 4 é uma condição necessária, mas não suficiente, para um U-invertido: em termos da expressão (7), $\beta_1 > 0$ e $\beta_2 < 0$, com significância. Isto é, temos evidências de que a derivada primeira de gasto em P&D contra IL é positiva, e a derivada segunda é negativa. Mas isso seria também compatível com relações diferentes do U-invertido, como por exemplo uma função côncava porém monótona. A fim de descartar essa última possibilidade, e verificar a existência do U-invertido, consideramos (e aplicamos aos nossos dados e estimativas) dois testes ou procedimentos:

- 1) O modelo estimado, do tipo

$$gastopd_i = \beta_0 + \hat{\beta}_1 \cdot IL_i + \hat{\beta}_2 \cdot IL_i^2 + \varepsilon_i \quad (8)$$

implica um ponto de máximo $IL^* = -\frac{\hat{\beta}_1}{2 \cdot \hat{\beta}_2}$ (9)

Uma evidência de relação na forma de U-invertido seria então mostrar que IL^* encontra-se dentro do intervalo de dados para IL . Tomando a regressão OLS I da tabela 4, e aplicando (9), temos que $IL^* = 0.277$; já tomando a regressão *tobit* IV da tabela 4, e aplicando (9), temos que $IL^* = 0.305$. Para ambas as regressões, o intervalo de dados $[IL_{\min}, IL_{\max}]$ da nossa amostra é $[0, 0.5]$, de modo que temos algumas observações à direita de IL^* , onde a curva $gastopd = f(IL, IL^2)$ já é descendente, evidenciando portanto a existência de um U-invertido.

- 2) Lind e Mehlun (2010) questionam o primeiro procedimento citado anteriormente. Intuitivamente, o argumento deles (aplicado ao nosso contexto) é o seguinte: uma vez que IL^* é uma estimativa, um teste forte da relação em U-invertido requereria não apenas que IL^* , mas também que um intervalo de confiança em torno de IL^* estivesse dentro do intervalo de dados. Esse teste forte seria equivalente a um duplo teste sobre a derivada $\frac{\partial(gastopd)}{\partial IL}$: para valores baixos de IL , deveríamos encontrar uma derivada positiva, e para valores altos de IL , uma derivada negativa, no caso de uma função com formato de U-invertido. Por sua vez, a partir de (8), essa derivada é:

$$\frac{\partial(gastopd)}{\partial IL} = \hat{\beta}_1 + 2 \cdot \hat{\beta}_2 \cdot IL \quad (10)$$

Substituindo os valores $IL_{\min} = 0$ e $IL_{\max} = 0.5$ em (10) obtemos estimativas da derivada nos pontos extremos do intervalo de dados. Levando em conta o fato de a derivada em (10) ser uma variável aleatória cuja variância depende das variâncias de $\hat{\beta}_1$ e de $\hat{\beta}_2$, chegamos às seguintes estatísticas- t :

$$t_j = \frac{\hat{\beta}_1 + 2 \cdot \hat{\beta}_2 \cdot IL_j}{\sqrt{\text{var}(\hat{\beta}_1) + 4 \cdot (IL_j)^2 \cdot \text{var}(\hat{\beta}_2)}} \quad j = \{\min, \max\} \quad (11)$$

A condição para o U-invertido consiste então em:

$$t_{\min} > t_{\alpha} \quad \text{e} \quad t_{\max} < -t_{\alpha} \quad (12)$$

onde t_{α} é o valor crítico da distribuição t de Student a $\alpha = 5\%$ de significância e para o número apropriado de graus de liberdade. Pela tabela 4, OLS I, temos 112

(observações) – 2 (regressores) = 110 graus de liberdade, e as variâncias de $\hat{\beta}_1$ e de $\hat{\beta}_2$ podem facilmente ser recuperadas a partir das estatísticas-t correspondentes. Um valor crítico aproximado, para o nosso caso, é $t_{\alpha} = 1.98$. Calculando então (11), obtemos $t_{\min} = 2.74$ e $t_{\max} = -0.92$. Vemos então que (12) não é satisfeita, e portanto que OLS I não passa no teste de Lind e Mehlun (2010) para o U-invertido. O mesmo resultado foi obtido para a regressão *tobit* IV.

Concluimos então, a partir dos dois testes aplicados, que existe apenas uma evidência fraca de que a relação entre gasto em P&D e poder de mercado (medido pelo IL) tem a forma de U-invertido.

5.2 Endogeneidade

Talvez a principal crítica feita à teoria Estrutura-Condução-Desempenho (ECD) em OI seja o efeito retroativo que a condução e o desempenho podem ter na estrutura da indústria. No nosso contexto, isso equivale a dizer que o gasto em P&D pode afetar a estrutura e o poder de mercado de uma indústria. Em termos econométricos, isso torna a variável IL endógena. O fato de um regressor ser endógeno torna o estimador OLS viesado e, portanto, inválido. A solução para esse problema consiste em obter uma variável correlacionada com poder de mercado e independente de choques aleatórios no gasto em P&D. Não é trivial obter uma variável com essas características para cada firma ou setor.

Uma possível variável deste tipo é obtida usando-se a condição de maximização de lucro do monopólio:

$$\frac{p - c}{p} = \frac{1}{\varepsilon_D}$$

onde p é o preço, c o custo marginal e ε_D é a elasticidade-preço da demanda. O lado esquerdo da equação é o IL, que na condição de primeira ordem se iguala ao inverso da elasticidade da demanda. Uma condição equivalente, mas em geral usando a demanda residual da firma, pode ser derivada para diversos modelos de oligopólio (ver, por exemplo, TIROLE, 1988), que descrevem melhor os setores que estamos analisando.

Essa relação teórica nos diz que o inverso da elasticidade é um possível instrumento para o poder de mercado. O porquê de a elasticidade representar uma variável exógena em relação ao poder de mercado pode ser entendido através de um modelo de oligopólio *a la* Cournot: o grau de poder de mercado e a estratégia de competição de uma firma em um mercado são resultado da interação desta firma com as demais, e derivam diretamente do equilíbrio do jogo. Esse equilíbrio por sua

vez é uma função dos parâmetros de custo das firmas e da demanda do mercado. Portanto, a elasticidade da demanda é um dos parâmetros que determina o equilíbrio do mercado e a estratégia de competição das firmas. Firmas que competem em inovação gastam em P&D. Enquanto a elasticidade da demanda é um parâmetro exógeno, o poder de mercado e o gasto em P&D são endógenos ao equilíbrio.

Esse raciocínio é válido quando assumimos pouca variação na gama de produtos ofertados pela firma e estabilidade da elasticidade da demanda. Algo verdadeiro no curto e médio prazo. No longo prazo, o gasto em P&D pode se materializar em inovação alterando a cesta de bens ofertados pela firma, a substitutibilidade destes e suas elasticidades-preço da demanda. Em síntese, nosso instrumento é válido se assumirmos que o gasto presente em P&D não afeta a demanda atual pelo produto da firma. Essa é uma hipótese razoável, pois supõe-se que gasto em P&D é uma política com resultados de longo prazo.

A dificuldade óbvia de se usar o inverso da elasticidade como uma medida de poder de mercado é que se faz necessário para conhecer a elasticidade-preço da demanda de todos os mercados das firmas da nossa amostra, uma informação não disponível e difícil de ser obtida via técnicas convencionais. Em vez de tentar estimar estas elasticidades, nós usamos a elasticidade percebida pela firma, e declarada em uma questão da pesquisa IBMEC-Sensus.¹⁹ Sem dúvida, uma crítica pertinente a esse nosso instrumento é que ele é baseado numa “percepção subjetiva”, declarada numa *survey* e, portanto, sem a mesma objetividade, digamos, da variável receita líquida que também utilizamos nas nossas regressões, e que é publicada no balanço anual das empresas.

A tabela 5 mostra os resultados (de segundo estágio) do modelo estimado usando o inverso da elasticidade como um instrumento para o IL. Os coeficientes têm o sinal esperado na maior parte das especificações, mas perdem significância. Tanto os modelos lineares – 2 Semiparametric Least Squares (2SLS) – quanto os modelos IV *tobit*. Vale ressaltar que no primeiro estágio os modelos estimados mostraram significância do instrumento. Os resultados do primeiro estágio encontram-se no apêndice C.

Esse é um resultado importante, pois mostra que a relação em U-invertido descrita por Aghion *et al.* (2002) é endógena na base IBMEC-Sensus 2008. Anteriormente, vimos na tabela 4 (para os modelos OLS e *tobit*) que os coeficientes associados a IL e IL^2 têm os sinais esperados e são significantes; ademais, nossas estimativas passaram no primeiro procedimento ou teste descrito na subseção 5.1 (o ponto de máximo IL^* encontra-se dentro do intervalo de dados para IL) – daí podemos concluir que poder de mercado (medido pelo IL) e gasto em P&D guardam uma relação na forma de U-invertido. Porém, o fato de que esses resultados

19. Ver a questão Q.55 na seção 4.

perdem significância uma vez que estimamos com variáveis instrumentais (tabela 5) nos diz que a relação de U-invertido pode ser uma consequência de variáveis omitidas relacionadas com o gasto em P&D e com poder de mercado, não havendo uma relação de causalidade entre eles.

TABELA 5

Resultados: variável instrumental – inverso da elasticidade(Variável dependente: *gastopd*)

	2SLS			IV <i>tobit</i>		
	I	II	III	IV	V	VI
<i>lerner</i>	106 <i>0.34</i>	13 <i>0.18</i>	51 <i>0.76</i>	153 <i>0.24</i>	758 <i>0.18</i>	212 <i>0.16</i>
<i>lerner</i> ²	-85 <i>-0.11</i>	60 <i>0.28</i>	-87 <i>-0.53</i>	-181 <i>-0.11</i>	-1650 <i>-0.16</i>	-326 <i>-0.09</i>
<i>dagro</i>	4.02 <i>0.64</i>			1.68 <i>0.13</i>		
<i>dcomerc</i>	6.68 <i>0.70</i>			-0.23 <i>-0.01</i>		
<i>dserv</i>	3.41 <i>1.34</i>			-1.87 <i>-0.39</i>		
<i>rec_liq</i>	2.04 <i>0.14</i>	5.77 <i>0.83</i>	3.60 <i>0.53</i>	7.87 <i>0.27</i>	-22.63 <i>-0.11</i>	3.83 <i>0.33</i>
<i>commodity</i>	-1.83 <i>-0.92</i>	-2.31 <i>-1.26</i>		-4.31 <i>-1.10</i>	-4.13 <i>-1.07</i>	
<i>rec_export</i>	-0.09 <i>-0.74</i>	-0.07 <i>-1.13</i>		-0.06 <i>-0.23</i>	0.07 <i>0.09</i>	
<i>_cons</i>	-4.22 <i>-0.36</i>	3.70 <i>0.97</i>	-0.98 <i>-0.42</i>	-8.33 <i>-0.35</i>	-28.70 <i>-0.21</i>	-19.20 <i>-0.44</i>
obs	100	100	100	100	100	154

Fonte: Elaboração dos autores.

Notas: Os números em itálico são estatísticas-t Student.

*significativo a 15%; **significativo a 10%; ***significativo a 5%.

6 CONCLUSÃO E DIREÇÕES FUTURAS DE PESQUISA

Este artigo procurou analisar a relação entre inovação (gasto em P&D) e poder de mercado. Primeiramente olhamos o que a teoria nos diz a este respeito. Aghion *et al.* (2002) mostram que a relação entre inovação e poder de mercado segue a forma de U-invertido: firmas operando em mercados muito competitivos ou em mercados com alto grau de poder de mercado gastam menos em P&D do que firmas operando em mercados com níveis intermediários de poder de mercado.

Em seguida, utilizamos a base IBMEC-Sensus 2008 para testar o resultado do U-invertido com dados brasileiros. A principal dificuldade aqui consiste em

determinar uma medida confiável de poder de mercado. Enquanto outros trabalhos usando dados brasileiros utilizaram o IHH, uma medida de concentração – e uma aproximação do poder de mercado – nós usamos o IL, uma medida mais apropriada de poder de mercado.

Nós encontramos uma correlação positiva (por volta de 30%) entre o IHH e o IL. Apesar desta correlação, os resultados econométricos se mostraram bastante distintos entre as duas medidas. Não obtivemos significância estatística com o IHH, mas sim com o IL – o que está de acordo com a teoria, que diz que poder de mercado, e não concentração, determina o gasto em P&D.

A nossa base de dados também nos permitiu abordar um outro problema desta literatura: o fato de que P&D também pode determinar a estrutura do mercado. Ou seja, poder de mercado é uma variável possivelmente endógena no nosso modelo. Para lidar com este problema, usamos a condição maximização de lucro de mercados oligopolistas, que diz que o *mark up* é proporcional ao inverso da elasticidade-preço da demanda. Também usamos a elasticidade declarada pelos respondentes da pesquisa como variável instrumental para o IL. Os resultados continuam com o sinal esperado, mas perdem significância estatística. Concluímos então que existe uma relação na forma de U-invertido entre P&D e poder de mercado, conforme obtido nos modelos que usam o IL como medida, e no sinal das variáveis de todos os modelos estimados, mas que essa pode não ser uma relação de causalidade, mas simplesmente uma consequência da correlação de P&D e poder de mercado com uma terceira variável omitida do modelo.

Futuramente, seria oportuno verificar se a relação em U-invertido entre gasto em P&D e poder de mercado (apropriadamente medido, tal como fazemos aqui, pelo IL) vale também para a base de dados da PINTEC.²⁰ Usando essa última base, seria também interessante estimar a relação entre *número de patentes* e poder de mercado, tal como fazem Aghion *et al.* (2002), e comparar os resultados obtidos com os da relação entre gasto em P&D e poder de mercado – conjecturamos se seria possível, a partir dessa comparação, recuperar evidências de retornos decrescentes na produção de patentes na indústria brasileira, tal como encontrado por Zachariadis (2002) para os Estados Unidos.

ABSTRACT

Since Schumpeter (1950), until more recent work such as Aghion et al. (2002), there exists a large literature analyzing the relationship between R&D and market power. This paper follows this tradition and analyzes the relationship between R&D expenditures and market power in the Brazilian economy. We use a new dataset: the IBMEC-Sensus 2008 survey. Besides information on R&D expenditure, this dataset contains

20. A fim de permitir um tratamento para o problema da endogeneidade, seria interessante que edições futuras da PINTEC incluíssem em seu questionário algo parecido com a questão da pesquisa IBMEC-Sensus que captava a elasticidade da demanda percebida pelas firmas.

an estimate of the price-elasticity of demand faced by the firm, information we use as an exogenous instrument for market power. We use the model from Aghion et al. (2002), which proposes an inverted-U relationship between R&D and market power – measured by the Lerner index – to justify the functional form we adopt in our econometric model. Our first results do not reject an inverted-U relationship between both variables. However, when we estimate the model using the inverse of the demand elasticity as an instrument for market power, this relationship loses significance, showing that the relation between R&D and market power could simply be a consequence of the simultaneous determination of both variables by a third and omitted variable. We conclude that there is no causality from market power to R&D.

REFERÊNCIAS

- AGHION, P.; HOWITT, P. A model of growth through creative destruction. **Econometrica**, v. 60, n. 2, Mar. 1992.
- _____. *et al.* **Competition and innovation**: an inverted U relationship. Institute for Fiscal Studies, 2002 (Working Papers, n. 02/04).
- BAUMOL, W. J. Contestable markets: an uprising in the theory of industry structure. **American Economic Review**, v. 72, n. 1, Mar. 1982.
- CGEE; ANPEI. **Os novos instrumentos de apoio à inovação**: uma avaliação inicial. Brasília, DF: Centro de Gestão e Estudos Estratégicos (CGEE) e Anpei, 2009.
- COHEN, W.; LEVIN, R.; MOWERY, D. **Firm size and r&d intensity**: a reexamination. NBER, 1987 (Working Papers, n. 2.205).
- GAZETA MERCANTIL. **Balanço Anual**, 2006.
- GROSSMAN, G.; HELPMAN, E. **Innovation and growth in the global economy**. MIT Press, 1991.
- JENSEN, J.; MENEZES-FILHO, N.; SBRAGIA, R. Os determinantes dos gastos em P&D no Brasil: uma análise com dados em painel. **Revista Estudos Econômicos**, v. 34, n. 4, 2004.
- KAMIEN, M. I.; SCHWARTZ, N. L. **Market structure and innovation**. Cambridge University Press, 1982.
- KANNEBLEY JUNIOR, S.; PORTO, G.; PAZELLO, E. **Characteristics of Brazilian innovative firms**: an empirical analysis based on PINTEC. 2005 (Research Policy, n. 34).
- LIND, J. T.; MEHLUN, H. With or without U? The appropriate test for a U-shaped relationship. **Oxford Bulletin of Economics**, v. 72, n. 1, 2010.
- PIRES-ALVES, C.; ROCHA, F. **Testing the Schumpeterian hypotheses for the Brazilian manufacturing industry**. Salvador: ANPEC, 2008.
- SCHERER, F. M. Market structure and the employment of scientists and engineers. **American Economic Review**, v. 57, n. 3, 1967.
- SCHMALENSEE, R. Inter industry studies of structure and performance. In: SCHMALENSEE, R.; WILLIG, R. D. (Ed.). **Handbook of Industrial Organization**, v. 2. chapter 16. Amsterdam: North Holland, 1989. p. 951-1.010.
- SCHUMPETER, J. A. **Capitalism, socialism, and democracy**. New York: Harper&Row, 1950.
- TINGVALL, P. G.; POLDAHL, A. Is there really an inverted U-shaped relation between competition and R&D? **Economics of Innovation and New Technology**, v. 15, issue 3, Mar. 2006.
- TIROLE, J. **The theory of industrial organization**. MIT Press, 1988.

ZACHARIADIS, M. **R&D, innovation, and technological progress**: a test of the Schumpeterian framework without scale effects. Louisiana State University, 2002.

(Originais submetidos em junho de 2009. Última versão recebida em março de 2011. Aprovada em abril de 2011.)

APÊNDICE A

QUESTÕES DA PESQUISA IBMEC-SENSUS

Q. 72) Gastos totais com P&D em 2006 (em R\$ correntes)

Comentário: Respondida por apenas 180 empresas, incluindo as que tiveram gasto = 0. Esta será a variável dependente dos nossos modelos OLS e *tobit*; a partir dela também, construímos a variável binária para os modelos *probit*.

Nome da variável de regressão: *gastopd*.

Q. 55) Em média, no ano de 2006, se você aumentasse o preço de seu principal produto em 10%, e tudo mais permanecesse alterado, a quantidade vendida: 1) teria uma queda de menos de 10%; 2) teria uma queda entre 10% e 30%; 3) teria uma queda entre 30% e 60%; 4) teria uma queda entre 60% e 100%; 5) não cairia; 6) Outras; e 99) Não sei.

Comentário: Esta é a questão a partir da qual pudemos inferir algo sobre a elasticidade da demanda. A fim de obter uma estimativa da elasticidade, convertemos as respostas/alternativas da Q.55 em: 1) 5%; 2) 15%; 3) 45%, e assim por diante. Depois, dividimos esses valores por 10% (variação do preço-proposta na questão).

Nome da variável de regressão: *invelast* (inverso da elasticidade).

Q. 59) Ainda considerando seu principal produto (carro-chefe), se tivesse de classificá-lo entre as duas categorias abaixo, você diria que ele está mais para: a) uma *commodity*, isto é, um produto já plenamente desenvolvido e absorvido pelo mercado, tal que só é possível expandir a quantidade vendida se houver redução de preço ou de custo; e b) um *bem diferenciado*, isto é, um produto que ainda pode sofrer significativos *quality upgradings*, mudanças no *design* e/ou na penetração no mercado, de tal forma que é possível no futuro expandir as vendas mesmo que o preço se mantenha constante ou até aumente.

Comentário: A partir desta questão podemos extrair uma variável binária captando oportunidade tecnológica, ou “tipo de inovação” (inovação de processo ou *catch-up* tecnológico *versus* inovação de produto).

Nome da variável de regressão: *commodity* (= 1 para *commodity*, = 0 para bem diferenciado).

Q. 62a) Qual a porcentagem das receitas do produto “carro-chefe” da sua empresa que foram geradas com exportações?

Comentário: Temos aí uma variável de controle para as nossas regressões, capturando o perfil exportador das empresas.

Nome da variável de regressão: *recexports*.

APÊNDICE B

TABELA B.1

IL e IHH a 3 dígitos de desagregação (correlação 0.22)

Lista de setores	IL	IHH	Lista de setores a 3 dígitos	Lerner	IHH
Aço	0.26	0.10	Fiação, tecelagem e confecções	-0.02	0.04
Açúcar	0.06	0.26	Franquias	0.02	0.23
Açúcar e álcool integrados	0.06	0.03	Frigoríficos	0.03	0.21
Administração de imóveis próprios	0.25	0.53	Frutas	0.04	0.10
Administradoras de rodovias	0.22	0.06	Gado bovino	-0.07	0.04
Adubos, fertilizantes e defensivos	0.02	0.15	Geração, transferência e transmissão de energia	0.07	0.15
Aerofotogrametria	0.14	0.24	Grandes fabricantes	-0.01	0.21
Agências de desenvolvimento urbano	-0.06	0.20	Hardware	0.07	0.45
Agências de viagem	0.15	0.18	Hospitais, clínicas e laboratórios	0.03	0.02
Água e saneamento	0.11	0.13	Hotéis	-0.01	0.06
Álcool	0.08	0.07	Iluminação	0.07	0.56
Alimentos à base de frutas	0.03	0.28	Imobiliárias e administração predial	0.39	0.06
Alimentos, bebidas e cereais	0.03	0.16	Incorporadoras de <i>shopping center</i>	0.26	0.11
Alumínio e artefatos	0.16	0.22	Incorporadoras e construtoras	0.03	0.05
Arroz, aveia, milho e féculas	0.04	0.13	Informática e telecomunicações	0.00	0.97
Artefatos de aço	0.14	0.29	Informática, papelarias e livrarias	0.04	0.27
Artefatos de madeira	0.08	0.15	Instalações elétricas, hidráulicas e de refrigeração	0.08	0.26
Artefatos de papel	0.06	0.08	Instrumentos de medição e precisão	0.08	0.59
Artefatos e utilidades de borracha	0.07	0.47	Internet	0.06	0.13
Artefatos e utilidades de plástico	0.06	0.07	Laticínios	0.03	0.15
Artigos de vestuário	0.05	0.07	Lentes e ótica	-0.04	0.52
Atacadistas de gás e combustível	0.02	0.31	Limpeza e segurança	0.02	0.18
Automação	0.01	0.24	Limpeza pública	0.04	0.18
Autopeças	0.06	0.15	Linha branca	0.00	0.37
Aves e ovos	0.07	0.32	Locação de veículos	0.14	0.51
Aviões, helicópteros e componentes	-0.01	0.95	Lojas de departamento e eletrodomésticos	0.02	0.11
Birôs de processamento	-0.01	0.11	Madeira e material elétrico e de construção	0.04	0.34
Bombas e compressores	0.08	0.50	Madeira para construção	-0.04	0.57
Brinquedos	0.02	0.24	Máquinas e equipamentos pesados	-0.03	0.21
Café	-0.09	0.27	Máquinas e ferramentas	0.05	0.29
Café, torrefações e solúvel	-0.05	0.18	Máquinas e implementos agrícolas	-0.05	0.05
Cal, gesso, cimento e artefatos	0.17	0.12	Máquinas industriais	0.07	0.08

(continua)

(continuação)

Lista de setores	IL	IHH	Lista de setores a 3 dígitos	Lerner	IHH
Calçados	0.08	0.13	Máquinas, ferramentas e ferragens	0.06	0.53
Caldeiraria	0.01	0.22	Máquinas-ferramenta	0.12	0.53
Cama, mesa e banho	-0.09	0.18	Material de construção, elétrico e de iluminação	-0.02	0.38
Caminhões e ônibus	0.04	0.05	Metais preciosos	0.37	0.66
Cana	0.06	0.17	Minerais não metálicos	0.17	0.44
Carroçarias	0.06	0.59	Moinhos, massas e pães	0.05	0.14
Carros, motos e utilitários	0.02	0.03	Montadoras	0.06	0.66
Carvão	0.05	0.34	Montagens industriais	0.06	0.16
Celulose e papel	0.10	0.12	Motores e componentes elétricos	0.15	0.72
Centrais de abastecimento	-0.06	0.39	Motores e componentes/retíficas	0.05	0.39
Cerâmicas, porcelanas e louças	-0.06	0.12	Móveis	0.44	0.23
Cereais e grãos	-0.03	0.12	Movimentação de pessoas e cargas	0.00	0.70
Cervejas	0.07	0.27	Não ferrosos	0.33	0.39
Chocolates, doces, balas e biscoitos	0.08	0.16	Navegação	0.09	0.30
Cigarros e fumo	0.18	0.72	Obras de infraestrutura	0.02	0.11
Cobre e artefatos	0.07	0.44	Óleos vegetais e animais	0.02	0.04
Cohabs	-2.00	0.05	Operadores logísticos	0.08	0.30
Comercialização	0.26	0.35	Outras bebidas	0.07	0.28
Componentes eletrônicos	0.05	0.46	Outros não ferrosos	0.14	0.19
Concretagem	-0.01	0.26	Partes, peças, acessórios e componentes automotivos	0.04	0.05
Condutores elétricos	0.10	0.24	Pavimentação e terraplanagem	0.08	0.12
Conglomerados alimentícios	0.02	0.21	Pedras, mármore e granitos	-0.03	0.21
Conservas de pescado	-0.02	0.84	Perfumarias	0.20	0.82
Conservas em geral	-0.01	0.35	Pesca	-0.03	0.19
Construção naval	0.04	0.48	Pesquisa e prospecção	-0.39	0.28
Construção pesada	0.08	0.13	Planos de saúde	0.03	0.04
Construtoras de conjuntos habitacionais	-0.47	0.37	Pneus	0.09	0.78
Consultorias	0.12	0.93	Porcas e parafusos	0.09	0.21
Cooperativas	0.07	0.09	Portos e aeroportos	-0.07	0.19
Culturas diversas	0.02	0.46	Produção e transmissão	0.11	0.17
Curtumes	0.03	0.11	Produtos agrícolas e alimentos	0.04	0.10
Derivados do petróleo	0.08	0.13	Produtos de higiene e limpeza	0.06	0.18
Desenvolvimento agrícola	-0.09	0.10	Produtos de papel	0.06	0.42
Destilados	0.02	0.45	Produtos farmacêuticos e veterinários	0.11	0.08
Distribuição	0.10	0.06	Produtos farmacêuticos, de higiene e limpeza	0.01	0.47
Distribuidores de aço	0.09	0.14	Produtos industrializados	0.03	0.21
Distribuidores de jornais e revistas	0.05	0.56	Produtos metalúrgicos	0.07	0.26
Diversos agricultura	-0.19	0.00	Produtos petroquímicos	0.05	0.13
Diversos alimentos	0.04	0.36	Produtos químicos e petroquímicos	0.00	0.06
Diversos autopeças	0.03	0.19	Produtos têxteis e de couro	-0.03	0.31
Diversos comércio atacadista	0.00	0.88	Projetos de arquitetura e engenharia	0.06	0.87
Diversos comércio exterior	0.03	0.20	Promoção de turismo	-0.41	0.06

(continua)

(continuação)

Lista de setores	IL	IHH	Lista de setores a 3 dígitos	Lerner	IHH
Diversos comércio varejista	0.03	0.14	Química fina	0.05	0.78
Diversos construção	0.10	0.09	Rações	0.01	0.18
Diversos couro e calçados	0.03	0.49	Recondicionamento de pneus	0.02	0.55
Diversos distribuidores de veículos e autopeças	0.08	0.48	Reflorestamento	0.01	0.14
Diversos eletrodomésticos	0.05	1.00	Refrigeração	0.20	0.50
Diversos energia elétrica	0.31	0.49	Refrigerantes e águas	0.12	0.16
Diversos equipamentos elétricos	0.13	0.57	Relógios e cronômetros	0.05	0.49
Diversos informática e tecnologia da informação	0.04	0.05	Relojoarias, joalherias e óticas	0.01	0.64
Diversos logística	0.08	0.84	Restaurantes e refeições prontas	0.02	0.33
Diversos madeira e móveis	0.16	0.97	Roupas e calçados	0.07	0.63
Diversos material eletrônico	-0.02	0.29	Sal	0.07	0.22
Diversos mecânica	0.08	0.12	Sementes e mudas	0.09	0.72
Diversos metalurgia	0.08	0.10	Serrarias	-0.03	0.35
Diversos mineração	0.05	0.24	Serviços especializados	0.19	0.09
Diversos minerais não metálicos	0.12	0.54	Serviços portuários e aeroportuários	0.11	0.32
Diversos papel e celulose	0.04	0.88	Serviços sociais	0.04	0.37
Diversos pecuária	0.00	0.15	Shopping center	4.83	1.00
Diversos plástico	-0.03	0.74	Software	0.11	0.13
Diversos química	0.11	0.12	Solos e fundações	0.08	0.15
Diversos saúde	0.02	0.14	Sucos	0.32	0.26
Diversos telecomunicações	0.21	0.18	Suínos	0.16	0.38
Diversos têxtil	0.06	0.10	Supermercados	0.00	0.38
Diversos turismo e alimentação	-0.02	1.00	Telefonia fixa e de longa distância	0.09	0.22
Edificações comerciais e residenciais	0.08	0.07	Telefonia móvel	-0.06	0.04
Eletrodomésticos	0.14	0.68	Telemarketing e call-center	0.05	0.27
Eletroportáteis	0.06	0.15	Temperos e condimentos	0.01	0.86
Embalagens	0.11	0.22	Terminais de carga e armazéns	0.12	0.05
Engenharia de tráfego/gerência de transportes	-0.08	0.67	Terminais de passageiros	0.00	0.41
Equipamentos de telecomunicações	0.03	0.28	Tintas, vernizes e solventes	0.05	0.18
Equipamentos médico-dentários	0.12	0.24	Transmissão de dados	0.12	0.54
Esportes e entretenimento	-0.12	0.06	Transporte aéreo	-0.07	0.18
Estanho e artefatos	0.06	0.88	Transporte de carga	-0.26	0.06
Estruturas metálicas para construção civil	0.08	0.14	Transporte de cargas especiais	0.29	0.33
Extração e refino	0.27	0.97	Transporte de passageiros urbano/rodoviário	-0.02	0.03
Farmácias e perfumarias	0.01	0.23	Tratores, implementos e máquinas rodoviárias	0.03	0.23
Ferragens e ferramentas	0.00	0.11	Trens e componentes	0.10	0.93
Ferro	0.11	0.07	Válvulas e conexões	0.05	0.38
Ferrosos	0.32	0.44	Varejistas de gás e combustível	0.10	0.13
Ferrovias e metrô	-0.05	0.11	Vidros e cristais	0.12	0.27
			Vinhos	0.03	1.00

Fonte: *Gazeta Mercantil* (2006).

APÊNDICE C

RESULTADOS DO PRIMEIRO ESTÁGIO

Observação: Variável dependente = *lerner*; variáveis explicativas = *inelast* + *controles*.

1) Modelo 2SLS com todas as variáveis

ivregress 2sls gastopd (lerner3dig2006 lerner2= inelastval inelastval2) commodity rec_liq recexports dagro dcomerc dserv, first

First-stage regressions

Number of obs = 100

F (8, 91) = 4.38

Prob > F = 0.0002

R-squared = 0.2778

Adj R-squared = 0.2144

Root MSE = 0.0892

Lerner	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
commodity	-0.0088014	0.0186008	-0.47	0.637	-0.0457497 0.0281469
recliqmilhao	0.0387107	0.0606297	0.64	0.525	-0.0817228 0.1591443
recexports	0.0015016	0.0004152	3.62	0.000	0.0006768 0.0023264
dagro	-0.0671438	0.0366	-1.83	0.070	-0.1398453 0.0055576
dcomerc	-0.0916965	0.0289741	-3.16	0.002	-0.14925 -0.0341431
dserv	-0.0243694	0.0242944	-1.00	0.318	-0.0726273 0.0238885
inelastval	-0.2063424	0.0927421	-2.22	0.029	-0.3905632 -0.0221216
inelastval2	0.0197109	0.0089051	2.21	0.029	0.002022 0.0373997
_cons	0.1900557	0.0514342	3.70	0.000	0.087888 0.2922233

Number of obs = 100

F (8, 91) = 3.27

Prob > F = 0.0025

R-squared = 0.2235

Adj R-squared = 0.1553

Root MSE = 0.0419

lerner2	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
commodity	-0.004148	0.0087336	-0.47	0.636	-0.0214962 0.0132003
recliqmilhao	-0.0025782	0.0284674	-0.09	0.928	-0.0591252 0.0539688
recexports	0.0007191	0.000195	3.69	0.000	0.0003319 0.0011064
dagro	-0.0324704	0.0171848	-1.89	0.062	-0.0666059 0.001665
dcomerc	-0.0243337	0.0136042	-1.79	0.077	-0.0513567 0.0026894
dserv	-0.0085289	0.0114069	-0.75	0.457	-0.0311874 0.0141296
inelastval	-0.0784988	0.0435451	-1.80	0.075	-0.1649957 0.0079982
inelastval2	0.0074731	0.0041812	1.79	0.077	-0.0008324 0.0157785
_cons	0.0607814	0.0241498	2.52	0.014	0.0128107 0.108752

2) Modelo *tobit* com todas as variáveis

ivtobit gastopdmilhao (lerner3dig2006 lerner2= invelastval invelastval2) commodity recliqmilhao recexports dagro dcomerc dserv, twostep ll first

First stage regressions

Source	SS	df	MS
Model	0.278767209	8	0.034845901
Residual	0.724573063	91	0.007962341
Total	1.00334027	99	0.01013475

Number of obs = 100
 F (8, 91) = 4.38
 Prob > F = 0.0002
 R-squared = 0.2778
 Adj R-squared = 0.2144
 Root MSE = 0.08923

Lerner	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
invelastval	-0.2063424	0.0927421	-2.22	0.029	-0.3905632 -0.0221216
invelastval2	0.0197109	0.0089051	2.21	0.029	0.002022 0.0373997
commodity	-0.0088014	0.0186008	-0.47	0.637	-0.0457497 0.0281469
recliqmilhao	0.0387107	0.0606297	0.64	0.525	-0.0817228 0.1591443
recexports	0.0015016	0.0004152	3.62	0.000	0.0006768 0.0023264
dagro	-0.0671438	0.0366	-1.83	0.070	-0.1398453 0.0055576
dcomerc	-0.0916965	0.0289741	-3.16	0.002	-0.14925 -0.0341431
dserv	-0.0243694	0.0242944	-1.00	0.318	-0.0726273 0.0238885
_cons	0.1900557	0.0514342	3.70	0.000	0.087888 0.2922233

Source	SS	df	MS
Model	0.045986006	8	0.005748251
Residual	0.159737315	91	0.001755355
Total	0.20572332	99	0.002078013

Number of obs = 100
 F (8, 91) = 3.27
 Prob > F = 0.0025
 R-squared = 0.2235
 Adj R-squared = 0.1553
 Root MSE = 0.0419

lerner2	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
invelastval	-0.0784988	0.0435451	-1.80	0.075	-0.1649957 0.0079982
invelastval2	0.0074731	0.0041812	1.79	0.077	-0.0008324 0.0157785
commodity	-0.004148	0.0087336	-0.47	0.636	-0.0214962 0.0132003
recliqmilhao	-0.0025782	0.0284674	-0.09	0.928	-0.0591252 0.0539688
recexports	0.0007191	0.000195	3.69	0.000	0.0003319 0.0011064
dagro	-0.0324704	0.0171848	-1.89	0.062	-0.0666059 0.001665
dcomerc	-0.0243337	0.0136042	-1.79	0.077	-0.0513567 0.0026894
dserv	-0.0085289	0.0114069	-0.75	0.457	-0.0311874 0.0141296
_cons	0.0607814	0.0241498	2.52	0.014	0.0128107 0.108752

UM MODELO ECONOMÉTRICO COM PARÂMETROS VARIÁVEIS PARA A CARGA TRIBUTÁRIA BRUTA BRASILEIRA TRIMESTRAL

Mario Jorge Cardoso de Mendonça*

Adolfo Sachsida*

Luis Alberto Medrano**

Este estudo tem por objetivo estimar um modelo econométrico linear com parâmetros variáveis para a análise da carga tributária bruta brasileira (CTBB) trimestral no período 1995-2009. A utilização desse tipo de modelo é justificada pelo fato de o sistema tributário nacional ter sido submetido a sucessivas mudanças durante boa parte do período em questão. Tais alterações ocorreram mormente em alíquotas e bases de incidência de tributos preexistentes. Ocasionalmente, somaram-se a elas eliminações (temporárias ou não) de alguns tributos e a criação de outros. As principais conclusões deste estudo são: *i)* o Produto Interno Bruto (PIB) é certamente a principal variável explicativa da dinâmica da CTBB no período em questão; *ii)* a elasticidade-PIB da CTBB parece estar próxima da unidade ou um pouco abaixo disso para duas medidas do PIB real utilizadas; *iii)* a parcela da CTBB autônoma em relação ao PIB teve mudança relevante durante todo este último período, possivelmente indicando contínuas melhorias nos procedimentos de fiscalização adotados pelas autoridades tributárias e/ou aumentos na formalização da economia; e, finalmente, *iv)* o modelo com parâmetro variando é o que melhor se ajusta ao problema aqui tratado para modelar o comportamento da CTBB.

1 INTRODUÇÃO

A carga tributária bruta (CTB) de uma determinada economia é dada pela soma dos recursos que o governo recolhe compulsoriamente do setor privado. Tudo mais permanecendo constante, aumentos na CTB reduzem a renda disponível do setor privado e, portanto, os recursos disponíveis para o financiamento das despesas de consumo e investimento de famílias e firmas. No entanto, incrementos na CTB aumentam (comumente, pelo menos) os recursos à disposição do governo e, por conseguinte, permitem diminuições no nível de endividamento público e aumentos na oferta de bens e serviços públicos à população. Assim, a dinâmica da CTB é um determinante crucial das dinâmicas de diversas outras grandezas macroeconômicas relevantes.

Infelizmente, a modelagem econométrica da CTB brasileira (CTBB) e de suas componentes é muito dificultada pelas constantes mudanças em alíquotas e bases de incidência de tributos preexistentes. Somem-se a isto ocasionais cessamentos e suspensões de alguns tributos e a criação de outros. Tais mudanças têm caracterizado a administração tributária do país no período pós-Plano Real. Não surpreende,

* Técnico de Planejamento e Pesquisa da Diretoria de Estudos e Políticas Macroeconômicas (Dimac) do Ipea.

** Pesquisador-bolsista do Programa de Pesquisa para o Desenvolvimento Nacional (PNPD) na Coordenação de Finanças Públicas (CFP) da Dimac/Ipea.

pois, que a literatura brasileira com especificações econométricas explícitas para a CTBB e seus principais elementos seja relativamente pequena e que os resultados relatados sejam frequentemente apenas exploratórios.

Este trabalho tem como objetivo contribuir para a literatura sobre o tema, apresentando um modelo econométrico linear com coeficientes variáveis para uma série trimestral da CTBB no período 1995-2008. A utilização de um modelo com coeficientes variáveis – discutido, entre outros, por West e Harrison (1997) – é justificada pela conjectura – que parece muito plausível *a priori* – de que as sucessivas mudanças no sistema tributário nacional estão associadas a mudanças recorrentes também nas elasticidades relevantes, de modo a tornar contraindicado o uso de especificações com parâmetros fixos.

A utilização de modelos com coeficientes variáveis para a modelagem de séries tributárias não é inédita na literatura brasileira (ver, por exemplo, HERNANDEZ, 1998; PORTUGAL; PORTUGAL, 2001). Mas, até onde se sabe, estes modelos jamais foram aplicados à série trimestral da CTBB.¹ Outra inovação do estudo, acredita-se, é a opção por modelar o componente sazonal da série da CTBB de forma endógena específica com base na representação de Fourier (WEST; HARRISON, 1997). Decidiu-se, ainda, estimar o modelo a partir de procedimentos bayesianos (GAMERMAN; LOPES, 1997; KIM; NELSON, 1999; WEST; HARRISON, 1997), que, dentre outras vantagens, permitem a obtenção de estimativas mais confiáveis em pequenas amostras do que métodos frequentistas (GELMAN *et al.*, 2003, p. 696).

Tendo em vista as considerações de Schettini *et al.* (2010), que apontam para a possibilidade de distorções associadas ao uso de diferentes índices de preço para deflacionar as variáveis econômicas, estimamos especificações para duas medidas distintas do Produto Interno Bruto (PIB) real. Na primeira delas, tomamos o índice do PIB real do Instituto Brasileiro de Geografia e Estatística (IBGE), (INDPIB), e na segunda usamos o PIB deflacionado pelo Índice Nacional de Preços ao Consumidor Amplo (IPCA), aplicando, assim, o mesmo deflator usado no cálculo da carga tributária real.²

Antecipando as conclusões, primeiro, observa-se que, ao longo do período estudado, o comportamento do PIB parece ter sido o principal determinante da dinâmica da CTBB. Segundo, os resultados mostraram que os valores obtidos da

1. Portugal e Portugal (2001) também trabalham com dados trimestrais, mas referentes apenas à arrecadação do Tesouro.

2. Tais diferenças entre o IPCA e o deflator implícito do PIB ocorrem principalmente pois: *i)* existem diferenças nas estruturas de ponderação dos índices e no peso das regiões; *ii)* a atualização das ponderações do deflator implícito do PIB ocorre anualmente, ao passo que tal ponderação é fixa no caso do IPCA – só é alterada em caso de uma nova Pesquisa de Orçamento Familiar (POF) –; e *iii)* os ajustes de consistência realizados pelo IBGE podem modificar tanto os valores nominais quanto os volumes do consumo final das famílias, alterando assim o deflator implícito.

elasticidade-PIB da CTBB para as duas medidas do PIB real foram muito próximas. Contudo, o resultado obtido com o índice do IBGE se situou em torno da unidade enquanto a elasticidade obtida com a variável PIB se mostrou um pouco menos elástica, ficando em torno de 0.90. Terceiro, diferentemente do que se supunha, não existe indicação de que esta elasticidade tenha sofrido variação acentuada ao longo do período. Em quarto lugar, pôde-se verificar que a parcela da CTBB autônoma em relação ao PIB tem se alterado durante o período. Por fim, verificou-se, após um exaustivo estudo, que o modelo onde tanto o termo autônomo como aquele referente à elasticidade do PIB é o que alcança a melhor especificação para modelar a CTBB.

O crescimento da parcela autônoma da CTBB pode ser uma indicação de contínuas melhorias nos procedimentos de fiscalização adotados pelas autoridades tributárias e/ou do crescimento exógeno verificado no grau de formalização da economia. Com efeito, grande parte dos modelos que consideram o termo autônomo da equação de regressão um parâmetro fixo termina por concluir que o valor da elasticidade-PIB da CTBB é significativamente maior que 1 (SANTOS; RIBEIRO; GOBETTI, 2008). Interpreta-se que este resultado se deve ao fato de que modelos com coeficientes fixos “forçam todo o peso” do crescimento recente da CTBB no PIB, subestimando (ou mesmo ignorando) melhorias na fiscalização tributária e incrementos no grau de formalização da economia.

Além desta introdução, o texto é constituído de outras quatro partes. A seção 2 apresenta alguns fatos estilizados acerca do tamanho e da composição da CTBB no ano de 2006. Na terceira seção, apresentam-se uma especificação econométrica e as motivações teóricas desta, bem como a descrição dos dados utilizados. Detalhes sobre a metodologia econométrica utilizada foram acrescentados em anexos. Os resultados obtidos por meio desta especificação – como várias outras especificações alternativas – são apresentados e analisados na quarta seção. Por fim, a seção 5 traz breves notas à guisa de conclusão.

2 EM QUE CONSISTE A CTBB?

A literatura brasileira é rica em metodologias alternativas de cálculo da CTBB. Neste estudo, será utilizada a metodologia das Contas Nacionais do IBGE, que tem as vantagens de: *i*) ser compatível com as diretrizes das Nações Unidas; *ii*) ser relativamente bem documentada; *iii*) ter se mantido razoavelmente constante nos últimos 15 anos; e, portanto, *iv*) ser comparável internacionalmente. De acordo com o IBGE (2008a, tabela sinótica 20), os tributos brasileiros podem ser divididos em cinco categorias: *i*) impostos sobre produtos; *ii*) outros impostos ligados à produção; *iii*) impostos sobre a renda e o patrimônio; *iv*) contribuições sociais efetivas; e *v*) impostos sobre o capital (tabela 1). A chamada CTB é simplesmente a soma dos itens *i* a *v*, dividida pelo PIB em um dado período contábil (tabela 1).

Em 2006, a CTBB de R\$ 808,6 bilhões representou 34,1% do PIB daquele ano, que atingiu R\$ 2.369,8 bilhões.

TABELA 1
Componentes da CTBB em 2006
(Em R\$ bilhões)

	Total	Federal	Estadual	Municipal
Total	808,6	541,6	219,1	47,8
Impostos sobre produtos	336,6	146,6	171,9	18,0
Outros impostos ligados à produção	30,5	16,3	7,9	6,3
Impostos sobre a renda e a propriedade	220,9	181,8	22,5	16,6
Contribuições previdenciárias (INSS ¹ + RPPS ² + FGTS ³ + PIS/PASEP ⁴)	216,8	196,9	15,7	4,2
Impostos sobre o capital	3,8	0,0	1,1	2,7

Fonte: IBGE (2008a, tabela sinótica 20).

Notas: ¹ Instituto Nacional do Seguro Social (INSS).

² Regime Próprio de Previdência Social (RPPS).

³ Fundo de Garantia do Tempo de Serviço (FGTS).

⁴ Programa de Integração Social/Programa de Formação do Patrimônio do Servidor Público (PIS/PASEP).

O restante desta seção discute o tamanho e a composição desses grandes grupos de tributos. São dois os propósitos deste esforço de recuperação histórica. Em primeiro lugar, sublinhar o fato de a CTBB ser um agregado complexo, formado por dezenas de tributos com lógicas e histórias diferentes. Em segundo lugar, enfatizar que as sucessivas mudanças na legislação tributária implicaram transformações muito significativas na composição da CTBB. Consequentemente, influenciaram na resposta relativa desta última variável a, por exemplo, mudanças no PIB ao longo do período em análise. Ambos os pontos são insumos cruciais para a correta interpretação dos resultados apresentados nas seções seguintes.

2.1 Impostos sobre produtos

São sete os principais impostos sobre produtos: *i*) Imposto sobre Circulação de Mercadorias e Prestação de Serviços (ICMS) (estadual); *ii*) Contribuição para o Financiamento da Seguridade Social (Cofins); *iii*) Imposto sobre Produto Industrializado (IPI); *iv*) Imposto sobre Serviços de Qualquer Natureza (ISS) (municipal); *v*) Imposto sobre Importações (II); *vi*) Imposto sobre Operações Financeiras (IOF); e *vii*) Contribuição de Intervenção no Domínio Econômico (Cide). Apenas o ICMS e a Cofins, somados, responderam por mais de três quartos de todos os impostos sobre produtos e por pouco menos de um terço da CTBB total em 2006. Ademais, os fatos geradores e a incidência destes dois tributos, e mesmo do ISS, podem ser tratados, em uma primeira aproximação, como iguais. Com efeito, ambos os tributos incidem (em boa parte pelo menos) sobre o valor agregado das empresas em geral. Logo, parece razoável supor que a arrecadação destes tributos

segue mais ou menos de perto a dinâmica do PIB. Note-se, entretanto, que, ao contrário do que ocorreu com o ICMS (estadual), a legislação da Cofins mudou consideravelmente no período 1995-2008. Houve aumento de alíquotas em 1999 e novamente em 2003-2004, desta feita em conjunto com uma radical mudança na sistemática de recolhimento da Cofins, com a redução da tributação sobre o faturamento e a criação de um regime de incidência não cumulativa.

Voltando-se a atenção para os demais impostos sobre produtos, verifica-se inicialmente que é natural supor que os IIs variem de acordo com o valor total destas últimas (que, por sua vez, é positivamente correlacionado ao PIB). Mas devemos lembrar que os IIs sofrem alterações também em virtude de acordos comerciais com outros países, ou ainda dependendo de alguma política nacional de estímulo a determinado setor específico. Os demais impostos sobre produtos são mais idiossincráticos. A título de ilustração: a arrecadação do IPI não depende apenas do PIB industrial (que é positivamente correlacionado com o PIB total, diga-se), mas também de considerações relativas à política industrial, setorial e de emprego. As recentes isenções concedidas ao setor automobilístico e a “linha branca” de eletrodomésticos atestam esta afirmação. Situações parecidas ocorrem com o IOF e a Cide-combustíveis. A arrecadação do IOF não depende exclusivamente do volume de operações financeiras registradas na economia (que, uma vez mais, é positivamente correlacionado com o PIB total), mas também de considerações relacionadas à política monetária. Ademais, as alíquotas do IOF são majoradas em períodos nos quais a cobrança da Contribuição Provisória sobre Movimentação ou Transmissão de Valores e de Créditos e Direitos de Natureza Financeira (CPMF) é cancelada, tal como ocorreu no início de 1999 e no início de 2008. Igualmente, a arrecadação da Cide-combustíveis – criada apenas em 2002, em mais um esforço de ajuste fiscal – não depende somente do consumo de combustíveis e lubrificantes, também positivamente correlacionada com o PIB total. Ela se subordina, outrossim, à política do governo para os preços destes itens e da própria administração tributária das empresas produtoras, notadamente a Petrobras.

Em suma, a análise tributo a tributo dos impostos sobre produtos sugere fortemente que a arrecadação destes impostos é positivamente correlacionada com o PIB. Indica, ademais, que dificilmente a elasticidade-PIB da arrecadação dos impostos sobre produtos permaneceu constante durante o período 1995-2008, tendo em vista a criação da Cide-combustíveis (e as constantes mudanças em suas alíquotas e bases de incidência), do IOF, do IPI e da Cofins. No caso desta última, houve mudança inclusive no regime de tributação. Também não devemos nos esquecer ainda de que em vários momentos ao longo do período 1995-2008 ocorreram alterações importantes no ICMS, notadamente durante o período mais agudo de guerra fiscal entre os estados.

2.2 Contribuições previdenciárias e outros impostos ligados à produção

A maior parte das contribuições previdenciárias, tanto na iniciativa privada quanto no setor público, incide sobre os rendimentos do fator trabalho, quer dizer, sobre a folha de pagamentos das firmas e dos governos. Isto é verdade, ainda, para as contribuições para o FGTS, que tecnicamente não são receitas públicas, e sim uma poupança compulsória dos trabalhadores e, em virtude disto, propriedade deles. A inclusão do PIB como variável explicativa da arrecadação desses tributos somente pode ser justificada pela hipótese de tal variável estar positivamente correlacionada com a massa de salários.

Entretanto, nem todas as contribuições previdenciárias incidem sobre a folha de pagamentos. Exceções importantes são o Sistema Integrado de Imposto e Contribuições das Microempresas e das Empresas de Pequeno Porte (Simples) e as contribuições para o PIS/PASEP, que somados atingiram R\$ 31,4 bilhões, ou pouco menos de 15% do total das contribuições previdenciárias em 2006. A rigor, o Simples, instituído em 1999 e modificado sucessivas vezes desde então, é um regime tributário diferenciado, simplificado e favorecido, que permite a micro e pequenas empresas (MPEs) pagarem um único tributo – sobre o faturamento – em substituição a vários outros: Imposto de Renda Pessoa Jurídica (IRPJ), Cofins, contribuições previdenciárias patronais etc. Na verdade, o Simples engloba o IRPJ, a Contribuição Social sobre o Lucro Líquido (CSLL), a Cofins, o PIS, contribuições previdenciárias patronais, o IPI (no caso de contribuinte desse imposto), o ICMS e o ISS (estes dois últimos, desde que o estado e o município tenham celebrado convênio com a União). A receita do Simples classificada como contribuição previdenciária é somente uma parte, ainda que a maior parte, do total recolhido ao Simples, sendo que o percentual destinado a essa rubrica aumenta gradativamente com o aumento do faturamento anual da empresa. Assim, a receita do Simples é dividida entre os vários tributos que a compõem e não é classificada totalmente como receita previdenciária. As contribuições para o PIS/PASEP, por sua vez, são em boa parte cobradas sobre o valor adicionado das firmas, mas apenas desde 2003, quando foi alterado seu regime tributário, tal como ocorreu com a Cofins. São classificadas como previdenciárias unicamente porque servem para capitalizar fundos com algumas características (nominais, pelo menos) previdenciárias. A inclusão do PIB como variável explicativa destes tributos se deve, portanto, à hipótese de que ele é positivamente correlacionado tanto com o faturamento das MPEs quanto com o valor adicionado das médias e grandes empresas.

Os relativamente pequenos Outros Impostos Ligados à Produção (OILPs) se dividem em dois grandes grupos. O primeiro deles é formado por tributos que, tal como ocorre com grande parte das contribuições previdenciárias, incidem sobre a folha de pagamentos, notadamente as contribuições para o salário-educação e para

o Sistema S.³ Destarte, o que foi dito sobre as contribuições previdenciárias (com exceção do Simples e do PIS/PASEP) é válido para estes tributos, cuja arrecadação somada atingiu R\$ 13 bilhões em 2006. Os demais OILPs (cuja arrecadação somou R\$ 17,4 bilhões em 2006) são formados por taxas setoriais diversas (por exemplo, fiscalização de telecomunicações, vigilância sanitária, poder de polícia etc.) com pouca ou nenhuma relação com o PIB.

Em síntese, pode-se dizer que: *i*) a maior parte da arrecadação das contribuições previdenciárias responde positivamente ao PIB; *ii*) uma pequena parte delas é autônoma com relação ao PIB; e *iii*) as quebras estruturais ocorridas na legislação incidente sobre as contribuições previdenciárias foram relativamente menos importantes do que as ocorridas com os impostos sobre produtos.

2.3 Tributos sobre a renda, o patrimônio e o capital

São quatro os principais tributos brasileiros sobre a renda, o patrimônio e o capital (IRPCs): o imposto de renda (IR), a CSLL das pessoas jurídicas, o Imposto sobre a Propriedade de Veículos Automotores (IPVA – estadual) e o Imposto Predial e Territorial Urbano (IPTU – municipal).

Quanto à arrecadação do IR, ela é usualmente dividida em seis componentes, quais sejam, o IR da pessoa física; o IR da pessoa jurídica; o IR retido na fonte sobre os rendimentos do trabalho; o IR retido na fonte sobre ganhos de capital; o IR retido na fonte sobre remessas de divisas ao exterior; e os outros tipos de IR retido na fonte. Com efeito, a arrecadação do IR atingiu R\$ 136 bilhões em 2006, ou pouco menos de 60,5% da arrecadação total dos IRPCs naquele ano. Para efeito de comparação, cumpre registrar que a CSLL – um tributo em quase tudo similar ao IR da pessoa jurídica (com a importante diferença de ser destinado primordialmente ao financiamento da seguridade social) – arrecadou R\$ 26,7 bilhões (ou 11,9% do total dos IRPCs), o IPVA arrecadou R\$12,3 bilhões (ou 5,5% dos IRPCs) e o IPTU arrecadou R\$ 13,3 bilhões (ou cerca de 5,9% dos IRPCs) naquele mesmo ano.

Além desses tributos, vale destacar a agora extinta CPMF. Cobrada entre 1997 e 2007, com uma significativa mudança de alíquota em meados de 1999, arrecadava pouco mais de 1% do PIB em média e representava perto de 15% de toda a arrecadação dos IRPCs. Tanto a criação quanto a extinção da CPMF representaram importantes *quebras estruturais* na série da CTBB. Finalmente, devem-se citar os impostos sobre transmissões *causa mortis* ou *inter vivos* (estaduais), cuja arrecadação atingiu R\$ 3,7 bilhões em 2006, ou 1,6% dos IRPCs.

3. Esse sistema abrange, entre outras, as seguintes entidades: Serviço Nacional de Aprendizagem Rural (Senar); Serviço Nacional de Aprendizagem do Comércio (SENAC); Serviço Social do Comércio (SESC); Serviço Nacional de Aprendizagem do Cooperativismo (SESCOOP); Serviço Nacional de Aprendizagem Industrial (Senai); Serviço Social da Indústria (Sesi); Serviço Social de Transporte (SEST); Serviço Nacional de Aprendizagem do Transporte (SENAT); Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (Sebrae).

Conclui-se, assim, ser natural conceber que as arrecadações do IR e CSLL dependem positivamente do PIB, dada a correlação positiva entre o PIB e os lucros das firmas e a massa salarial da economia. O mesmo é verdadeiro no caso da CPME, enquanto foi cobrada. No entanto, parece lícito supor que as arrecadações (menores) do IPTU, IPVA e dos impostos sobre transmissões *causa mortis e inter vivos* sejam relativamente autônomas em relação ao PIB.

3 AS ESPECIFICAÇÕES ECONÔMETRICAS E OS DADOS UTILIZADOS

Claro está que a arrecadação de um determinado tributo depende positivamente tanto da alíquota quanto do montante do fato gerador. Dito de outro modo, é natural supor, à guisa de exemplo, que, tudo mais permanecendo constante, a arrecadação do II crescerá com o volume das importações, ou que a arrecadação do IOF crescerá com o volume das operações de crédito. Note-se, todavia, que o elevado número de tributos existentes no Brasil faz com que a estimação da CTBB tributo a tributo seja por demais trabalhosa, tornando comum a utilização do PIB como *proxy* para os fatos geradores de todos os tributos brasileiros (ver PORTUGAL; PORTUGAL, 2001; HERNANDEZ, 1998; SANTOS; RIBEIRO; GOBETTI, 2008). Naturalmente, espera-se que elevações e quedas no PIB real estejam associadas às elevações e quedas na arrecadação tributária real.

A inflação é outro fator usualmente listado como uma variável explicativa importante da dinâmica da CTB medida em termos reais. Seu sinal é incerto, contudo. Se, por um lado, elevações da inflação podem diminuir a arrecadação (real) caso os pagamentos (nominais) de tributos sejam recolhidos com atraso pelos contribuintes (TANZI, 1977), por outro, tais elevações reduzem o valor real das faixas de isenção (admitindo-se que elas fiquem fixas em níveis nominais), por exemplo, do IR. Com isso, alarga-se a base de arrecadação do imposto e, desta forma, a própria arrecadação tributária. Em resumo, não surpreende que a inflação afete as receitas públicas de modos diferentes em contextos diferentes.

Vários estudos (HAKKIO; RUSH, 1991; BOHN, 1991, 1998) buscam explicar a dinâmica dos fluxos das receitas e dos gastos correntes do governo a partir de considerações sobre as implicações dinâmicas destes fluxos sobre o endividamento público. A hipótese básica dos estudos é a de que os governos fixam suas receitas e gastos de modo a manter o endividamento público sob controle, por assim dizer. Parece válido, pois, acrescentar o comportamento da dívida líquida do setor público (DLSP) aos determinantes (potenciais, ao menos) da dinâmica da CTBB.

Tendo em vista as considerações apresentadas, optou-se por iniciar a investigação econométrica com a especificação do Modelo Linear Dinâmico – Dynamic Linear Model (DLM) – que aparece na equação (1). A introdução da “constante” nesta especificação não é ingênua, já que, como visto na seção 2, pode-se supor

que uma parcela da arrecadação tributária – aquela devida à tributação sobre o patrimônio, por exemplo –, não está relacionada a nenhum dos fatores descritos anteriormente.

$$\begin{aligned} CTBB_t &= b_0(t) + b_1(t)PIB_t + b_2(t)DLSP_t + b_3(t)INFLA_t + S_t + v_t \\ b_i(t) &= \varphi_i b_i(t-1) + w_i(t), \quad i = 0, \dots, 3 \end{aligned} \quad (1)$$

onde $v(t) \sim N(0, V)$, $w_i(t) \sim N(0, W_i)$, para $i = 0, \dots, 3$, enquanto φ_i é o componente autorregressivo da equação de estado para cada b_i . S_t é o componente sazonal, que é modelado de acordo com a representação de Fourier (WEST; HARRISON, 1997), de modo que:

$$S_t = \sum_{j=1}^2 \left[a_j \operatorname{sen}\left(\frac{2\pi jt}{4}\right) + b_j \cos\left(\frac{2\pi jt}{4}\right) \right] = Z_t \gamma$$

onde:

$$Z_t = \left(\operatorname{sen}\left(\frac{2\pi t}{4}\right) \operatorname{sen}\left(\frac{2\pi 2t}{4}\right) \cos\left(\frac{2\pi t}{4}\right) \cos\left(\frac{2\pi 2t}{4}\right) \right)^T$$

e:

$$\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$$

De modo a fazer com que as estimativas obtidas expressem os valores das elasticidades, empregou-se aqui a transformação logarítmica, excetuando-se a variável INFLACÃO. Os dados trimestrais se referem ao período de janeiro de 1995 a dezembro de 2008 e são mostrados graficamente no anexo A. Conforme pode ser visto, a CTBB apresenta um destacado padrão sazonal tanto em valores reais quanto em relação ao PIB. Daí a necessidade de incluir o componente sazonal supracitado no modelo. Tendo em vista o fato de que, devido às questões metodológicas, o índice do PIB real medido pelo IBGE é levemente inconsistente com a série da CTBB empregada neste estudo quando tomado como fato gerador desta última, iremos inicialmente usar o PIB em valor real como medida preferencial para o fato gerador. A seguir, o exercício será revisto tomando desta vez o índice calculado pelo IBGE. Uma vantagem disto é que podemos mais facilmente comparar nosso resultado

com os de outros estudos, tendo em vista que este indicador é o que comumente se emprega na literatura. Assim, as variáveis usadas neste estudo são as seguintes:

- CTBB: carga tributária bruta brasileira nominal deflacionada pelo IPCA medido pelo IBGE.
- DLSP: valor real da dívida líquida do setor público consolidado deflacionada pelo Índice Geral de Preços – Disponibilidade Interna (IGP-DI), calculada pelo Banco Central do Brasil (BCB).
- INFLACÃO: taxa de inflação calculada partir da variação da média do IPCA no trimestre.
- PIB: produto interno bruto medido a preços de mercado e deflacionado pelo IPCA.
- INDPIB: índice do produto interno bruto medido a preços de mercado calculado pelo IBGE.

Devemos ressaltar que o procedimento bayesiano adotado para a estimação do DLM, elimina o problema da ordem de integração das séries bem como diminui a relevância do tamanho da amostra. Diversos autores (DEJONG; WHITEMAN, 1991; KOOP, 1992; SIMS, 1988; SIMS; UHLIG, 1991) advogam vigorosamente em favor da alternativa bayesiana sobre a abordagem clássica mais tradicional usada no que se refere à questão da raiz unitária. Por exemplo, é conhecido que os testes Augmented Dickey Fuller (ADF) são de baixa potência diante das alternativas plausíveis, especialmente no que se refere à hipótese alternativa de tendência estacionária. A abordagem bayesiana, por outro lado, revelaria que as hipóteses de raiz unitária e de tendência estacionária apresentariam probabilidade bastante similares quanto às suas funções de distribuição *a posteriori*. Assim, a abordagem bayesiana fornece um sumário mais razoável da informação amostral que a abordagem clássica.

Um outro problema com os testes clássicos de raiz unitária é o da descontinuidade gerada na teoria assintótica (SIMS, 1988). A abordagem bayesiana desde que baseada na função de distribuição *a posteriori* não apresenta o problema da descontinuidade. Koop (1994) assinala que, na abordagem clássica, os valores críticos gerados a partir de pequenas amostras podem diferir substancialmente dos valores críticos assintóticos. A abordagem bayesiana, desde que condicional à amostra observada, fornece resultados para pequenas amostras mais exatas. Em resumo, o problema da raiz unitária não é um ponto crítico na estatística bayesiana.

Por fim, diferentemente da abordagem clássica, a prática bayesiana não é dependente do teorema central do limite, o que elimina o uso de propriedades assintóticas (GELMAN *et al.*, 2003, p. 696). Isso traz consigo uma importante vantagem em termos da habilidade de obter estimativas mais confiáveis num

contexto de pequenas amostras e alta dimensionalidade paramétrica. Este ponto, por sua vez, é fundamental no caso brasileiro, onde séries de variáveis macroeconômicas realmente confiáveis somente estão disponíveis a partir do aparecimento do Plano Real.

4 RESULTADOS

Nesta seção são apresentados os resultados da estimação do modelo inicial definido pela equação (1), assim como os resultados oriundos de especificações alternativas. A metodologia econométrica utilizada nas estimações é relativamente complexa. Uma descrição sucinta acerca da implementação do nosso modelo está descrita no anexo B.

Após um exaustivo estudo procurando identificar o melhor conjunto de variáveis capaz de modelar a CTBB, observamos que nem a inflação nem a DLSP acrescentam informação relevante ao modelo. Em todas as especificações onde estas variáveis foram utilizadas a mediana do intervalo de confiança (IC) se situou em torno de zero ao longo de todo o período, sendo somente o PIB se mostrado como variável significativa na equação da CTBB. Deste modo, considerou-se como melhor alternativa tomar como referência especificações em que somente o PIB ou o INDPIB aparece como variável explicativa.

Com o objetivo de deixar claras as implicações de se usar medidas alternativas do produto e de se permitir (ou não) que cada parâmetro varie no tempo, uma estratégia utilizada aqui e também de sorte a obter a melhor especificação foi testar um conjunto de cinco modelos: *i*) um modelo básico tendo a variável INDPIB e todos os parâmetros variando; *ii*) um modelo 2 que se diferencia do básico apenas pelo uso da variável PIB; *iii*) um modelo 3 que se diferencia do modelo básico apenas pela imposição de um parâmetro fixo para o INDPIB; *iv*) um modelo 4 que se diferencia do básico apenas pela imposição de um parâmetro fixo para o intercepto; e *v*) um modelo 5 que se distingue do modelo básico pela imposição de parâmetros fixos para o intercepto e o INDPIB.

A seleção do melhor modelo foi feita com base sobretudo nos critérios de erro quadrático médio (EQM) e desvio absoluto médio (DAM), que avaliam a projeção fora da amostra (neste caso, implementada quatro passos à frente). Em todos os modelos, a sazonalidade foi tratada a partir da aplicação da representação de Fourier conforme definida na seção 3.

As tabelas 2 e 3 mostram, respectivamente, os valores obtidos para os critérios EQM e DAM com relação ao grau de ajustamento do modelo implementado dentro da amostra, bem como aquele gerado a partir da previsão feita quatro passos à frente fora dela.

TABELA 2

Ajustamento do modelo

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
EQM	0,0012	0,0001	0,000615	0,0034494	0,02573504
DAM	0,0316	0,0334	0,07375	0,1683688	0,4929952

Fonte: Elaboração dos autores.

TABELA 3

Previsão quatro passos à frente fora da amostra

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
EQM	0.0001	0.0012	0.002603	0.029418	0.00403807
DAM	0.0334	0.0316	0.04753	0.107508	0.06056104

Fonte: Elaboração dos autores.

Em ambas as tabelas é perceptível que os modelos 1 e 2 são os que apresentam melhor desempenho segundo os critérios EQM e DAM. Tomando por base o grau de ajustamento do modelo, verifica-se que o modelo 2 é o mais ajustado segundo o EQM, enquanto isso ocorre no modelo 1, caso se observe o valor do DAM. No caso da previsão fora da amostra, o que se verifica é exatamente o contrário. Os ajustamentos dos modelos 3 e 4 mostraram ajustamento e previsão fora da amostra sensivelmente piores que nos modelos 1 e 2. De acordo com os resultados das tabelas 2 e 3 justifica-se o emprego da regressão com parâmetros variáveis para modelar a CTBB, pelo fato de os modelos 1 e 2 serem superiores qualquer que seja o critério utilizado quando comparados àqueles em que existam parâmetros fixos.

Os gráficos 1 e 2 mostram, respectivamente, a evolução do coeficiente estimado do INDPIB e do intercepto para o modelo 1. O gráfico 3, por sua vez, mostra a evolução do coeficiente do PIB para o modelo 2. Nestes gráficos são mostrados o IC a 95% assim como sua mediana. Pode ser visto nestes gráficos que a elasticidade da carga tributária para ambas as *proxies* de renda tem se mantido praticamente constante ao longo de todo o período. Contudo, de modo que o leitor possa melhor visualizar, apresentamos no anexo C os gráficos em que aparece apenas a mediana deste coeficiente. É interessante notar que mesmo acontecendo numa escala diminuta existe um crescimento ao longo do tempo deste parâmetro. Nota-se ainda que a elasticidade da CTBB em relação à variável PIB é menos elástica que aquela medida em relação ao INDPIB. Para a primeira temos valores inicial e final de 0,88 e 0,91, enquanto para a segunda temos para valores inicial e final, respectivamente, 0,95 e 1,01. A tabela 4 apresenta o IC para a elasticidade do INDPIB para o modelo 3. Por fim, os resultados para os demais modelos são apresentados no anexo D.

GRÁFICO 1

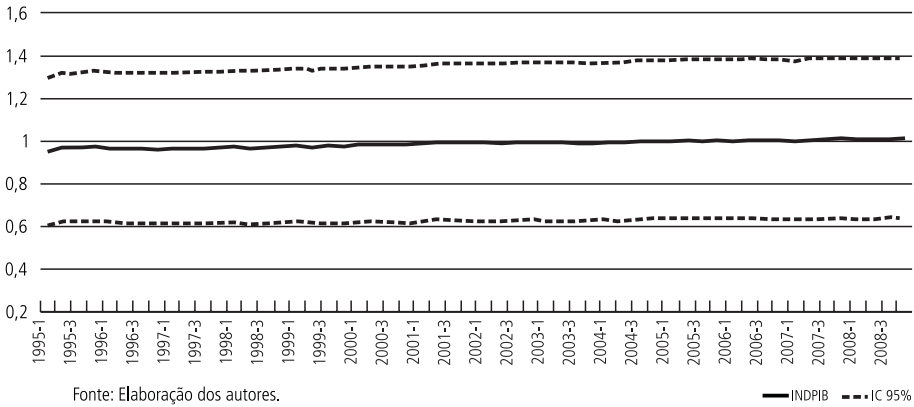
Mediana e IC 95% a posteriori de β (INDPIB)

GRÁFICO 2

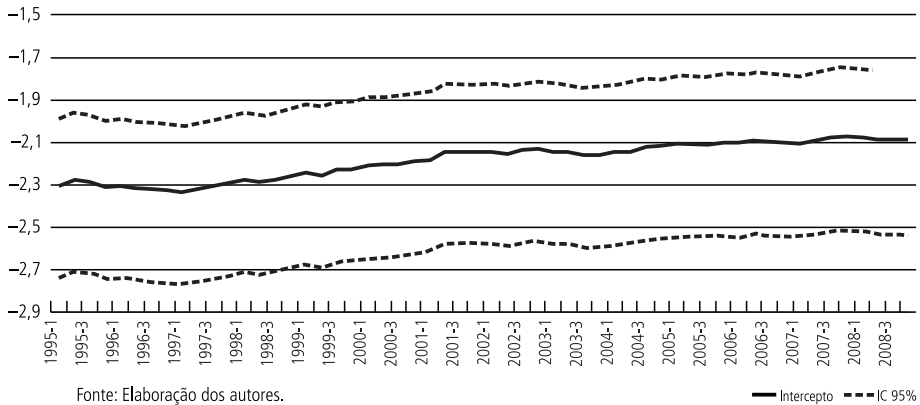
Mediana e IC a posteriori de β_0 (intercepto)

GRÁFICO 3

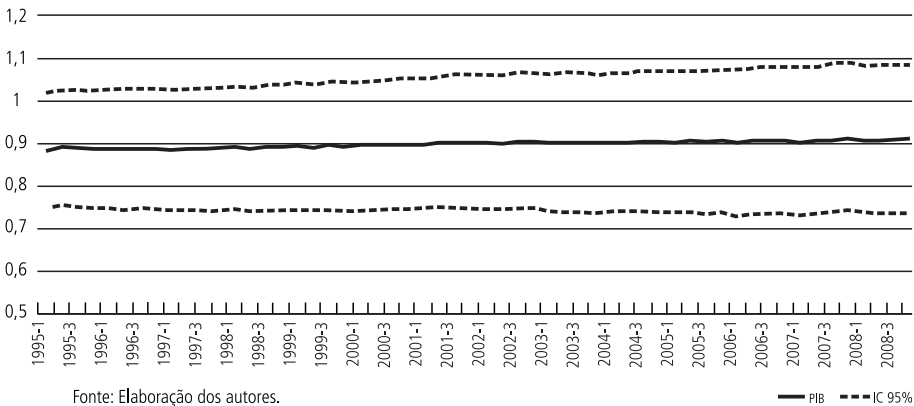
Mediana e IC a posteriori de β_1 (PIB)

TABELA 4
Estimativas INDPIB fixo (modelo 3)

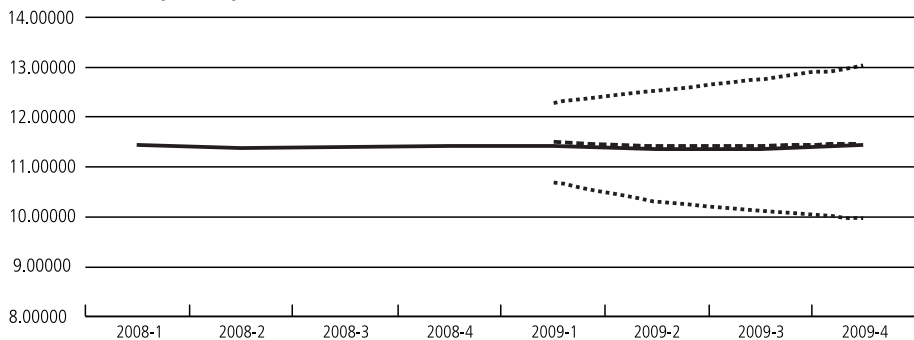
θ	E(θ)	IC 95%		R
Beta 1	0.9435	0.8612	1.060735	1,0002

Fonte: Elaboração dos autores.

Em relação ao intercepto, contudo, observa-se um crescimento mais firme. Neste caso, o valor médio inicial ficou próximo a 2,00, enquanto o valor médio final se situou em 2,40, conforme se observa no gráfico 1. Uma interpretação econômica para isto pode ser feita a partir do reconhecimento de que fatores não explicitamente incluídos no modelo estariam agindo para fomentar o crescimento da carga tributária. Esses fatores podem incluir a diminuição do grau de informalidade da economia, o crescimento do emprego formal – tal como reportado por Corseuil e Foguel (2009) – ou, ainda, indicar uma eficiência maior na arrecadação tributária (melhora dos procedimentos de fiscalização, por exemplo).

O gráfico 4 ilustra a previsão quatro passos à frente fora da amostra para o modelo 2, enquanto o ajustamento do mesmo modelo é demonstrado no gráfico 5. Neste gráfico, pode-se observar o alto grau de ajustamento. A tabela 5 replica o gráfico 4, mostrando os valores após os valores serem revertidos do log. O baixo erro de previsão em todos os casos revela a qualidade da previsão obtida.

GRÁFICO 4
Previsão para o período Q1/2009-Q4/2009



Fonte: Elaboração dos autores.

— CTB95 - - - Mediana IC 95% - . - . IC 95%

A tabela 6 mostra a média e o IC de 95% para a variância V da equação do modelo 2, os coeficientes da sazonalidade $\gamma_1, \gamma_2, \gamma_3$ e γ_4 , assim como as variâncias W_1, W_2 e W_3 e os coeficientes autorregressivos ϕ_1, ϕ_2 e ϕ_3 das equações que descrevem a variação dos parâmetros do modelo ao longo do tempo (seção 3). Por fim, a tabela 6 apresenta o resultado da estatística Gelman-Rubin R para se conferir a

convergência dos parâmetros gerados por meio do processo de estimação do modelo – o chamado algoritmo Monte Carlo Markov *Chains* (MCMC) (ver anexo B). A estatística R é obtida a partir de múltiplas cadeias markovianas obtidas com o referido processo de estimação. A convergência é alcançada quando o valor de R se situa próximo a 1, abaixo do nível crítico. Cadeias independentes foram geradas para 15 mil simulações. Em todos os casos as cadeias convergiram.

GRÁFICO 5
Ajustamento do modelo

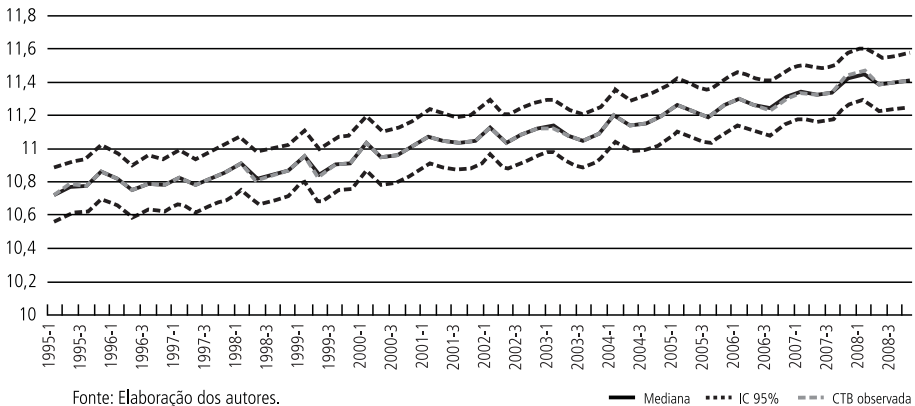


TABELA 5
Previsão quatro passos à frente
(Em R\$ milhões)

Período	CTBB	Previsão	Erro de previsão (%)
2009/01	91.680	97.792	6,67
2009/02	84.997	90.435	6,40
2009/03	84.948	91.069	7,49
2009/04	93.194	94.069	0,94

Fonte: Elaboração dos autores.

TABELA 6
Estimativas pontuais e ICs de 95% de para V , W , Γ e ϕ

θ	$E(\theta)$	IC 95%		R
V	0,010696	0,004717	0,020488	1,003945
W_1	0,014318	0,005224	0,031011	1,007219
W_2	0,001187	0,000821	0,001672	0,999788
γ_1	0,071133	-0,02387	0,168906	1,009629
γ_2	-0,00663	-1,63802	1,693714	0,999783
γ_3	0,007329	-0,09643	0,110495	1,005529
γ_4	-0,01976	-0,07697	0,033584	0,999875

Fonte: Elaboração dos autores.

5 CONSIDERAÇÕES FINAIS

Nas seções anteriores foi apresentado e contextualizado um modelo econométrico linear com parâmetros variáveis para a dinâmica da CTBB trimestral. Tanto no que toca ao ajustamento dentro da amostra, quanto no que se refere à qualidade das previsões do modelo fora da amostra, os resultados obtidos com o modelo foram muito satisfatórios. Em particular, o erro de previsão do modelo fora da amostra do valor total da CTBB em 2008 foi inferior a 8% para todos os períodos projetados para 2009.

Alguns estudos apresentam valores para a elasticidade da CTBB em relação ao PIB situadas em torno da unidade (HERNANDEZ, 1998; PORTUGAL; PORTUGAL, 2001). Outros estudos que consideram o termo autônomo da equação de regressão um parâmetro fixo terminam por concluir que o valor da elasticidade-PIB da CTBB é significativamente maior que 1 (SANTOS; RIBEIRO; GOBETTI, 2008). Neste estudo, estimamos o mesmo modelo usando duas *proxies* distintas para o fato gerador da CTBB. Os resultados indicam que a elasticidade da CTBB em relação ao produto real parece realmente ser unitária ou então próximo disso. A elasticidade obtida usando o índice do PIB medido pelo IBGE indicou o valor da unidade, enquanto este parâmetro cai para próximo de 0,90 quando a elasticidade da CTBB é tomada em relação ao PIB deflacionado pelo IPCA. Assim, os resultados do presente trabalho se aproximam daqueles encontrados em Hernandez (1998) e Portugal e Portugal (2001) se afastando um pouco daquilo obtido por Santos, Ribeiro e Gobetti (2008).⁴

A questão que se coloca é: por que os estudos citados acima encontram resultados distintos em relação à elasticidade do PIB no modelo da carga tributária? De modo a tratar esta questão, foram estimadas outras especificações do modelo tal como é apresentado no início da seção 4. Os resultados mostraram que o modelo mais bem especificado, segundo os critérios adotados, é o modelo com parâmetros variando no tempo tanto para a elasticidade da *proxy* da renda como para o termo autônomo. Tendo em vista este resultado, constatamos que ambos os parâmetros sofrem variação no tempo, embora isso seja mais perceptível em relação ao termo autônomo. O resultado obtido para a elasticidade da renda não sofre tanta distorção em relação ao modelo básico onde os parâmetros não variam no tempo quando permitimos mudança no intercepto, mas quando mantemos a elasticidade da renda como parâmetro fixo. Contudo, a alteração para o coeficiente do termo autônomo em relação ao modelo básico é acentuada quando se permite que este coeficiente varie no tempo.

4. Deve-se ter em mente que Hernandez (1998) e Portugal e Portugal (2001) usam uma amostra de dados cuja dimensão temporal e frequência são distintas daquela utilizada por nós, enquanto Santos, Ribeiro e Gobetti (2008) fazem uso da mesma base de dados.

O diagnóstico que encontramos é que métodos econométricos que não permitem variação contínua em todos os parâmetros do modelo não são indicados para estimação de um modelo da CTBB porque, como mencionado na introdução, a evolução da CTBB é um fenômeno que está sujeito a contínuas mudanças oriundas de vários fatores.

Por fim, deve-se atentar para o fato de que a previsão da CTBB foi feita tendo em vista o conhecimento prévio das variáveis exógenas no horizonte de previsão, o que não ocorre numa situação real em que se deseja fazer a previsão para o futuro. Neste caso, teríamos de prever as variáveis exógenas antes de incorporá-las ao modelo da CTBB, o que geraria mais uma fonte de erro. Porém, isso não invalida o estudo no sentido de se determinar quais são as principais variáveis condicionantes da CTBB bem como a elasticidade da carga tributária em relação ao PIB, parâmetro fundamental na elaboração da política macroeconômica.

ABSTRACT

This article estimates a linear econometric model with variable parameters for the analysis of the Brazilian quarterly gross tax burden in the period 1995-2008. The choice of this particular model was motivated by the changes happened in the Brazilian tax system during these years – most frequently in the scope, design, and size of pre-existing taxes, the elimination of some of these taxes and/or the creation of new ones. The main conclusions of the paper are that: *i*) GDP is, quite certainly, the main explanatory variable of the dynamics of the quarterly tax burden during the period in question; *ii*) the GDP-elasticity of the Brazilian tax burden appears to have fluctuated around values close to unity; *iii*) the fraction of the Brazilian tax burden that does not depend on GDP (or on any other variable, for that matter) increased quite significantly during the most recent years – possibly due to continuous improvements in the ability of the Brazilian government to collect tax revenues and/or the increase in the formal sector of the economy; and finally *iv*) the model with variable parameters is the most appropriate to deal with this problem considering the criteria adopted.

REFERÊNCIAS

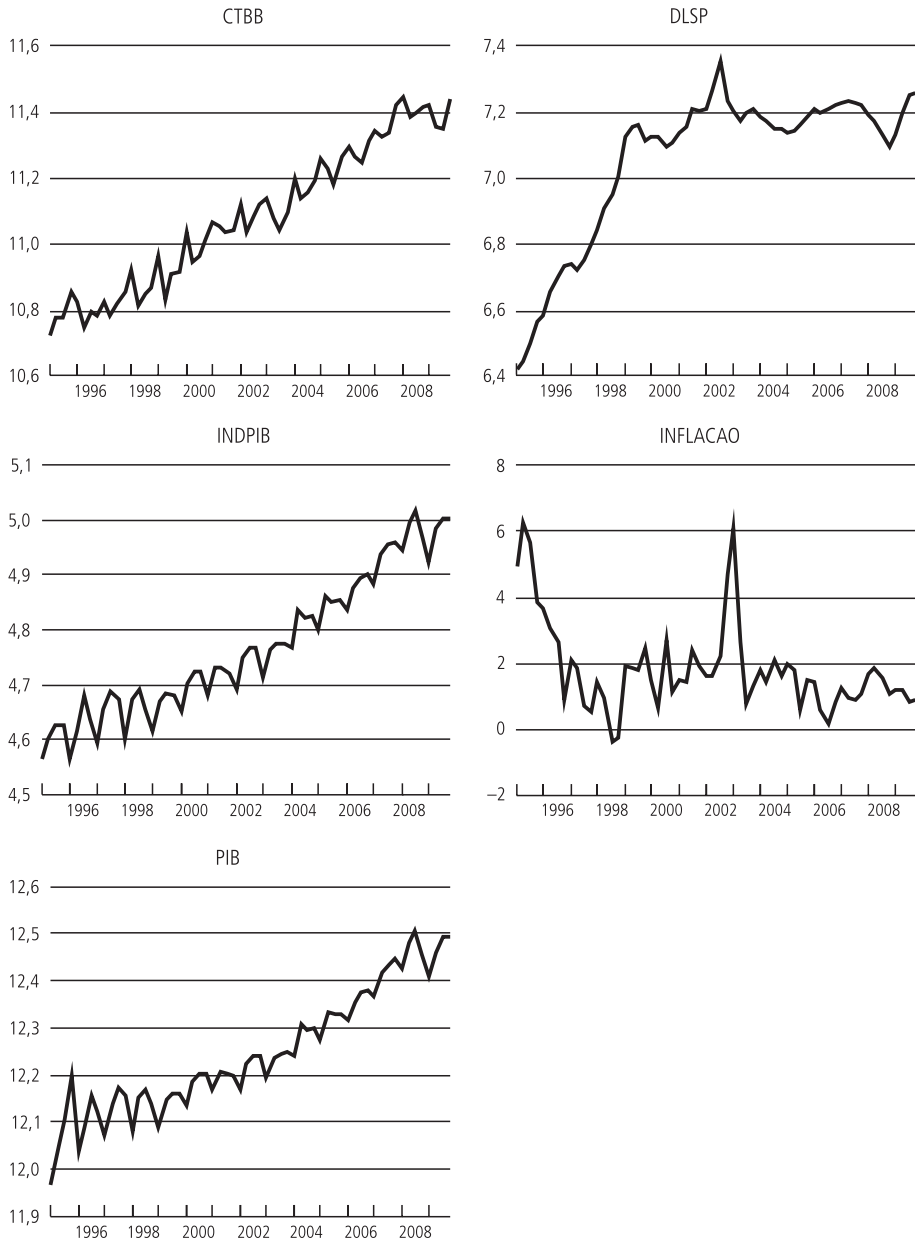
- BOHN, H. The sustainability of budget deficits with Lump-Sum and with income-based taxation. **Journal of Money, Credit and Banking**, v. 23, n. 3, p. 581-604, 1991.
- _____. The behavior of U.S. public debt and deficits. **The Quarterly Journal of Economics**, n. 113, p. 949-963, 1998.
- BRASIL. Ministério da Fazenda. Secretaria da Receita Federal. **Carga Tributária no Brasil 2008**. Brasília, 2009.
- CARTER, C. K.; KOHN, R. On gibbs sampling for state space models. **Biometrika**, v. 81, n. 3, p. 541-553, 1994.
- CORSEUIL, C. H. L.; FOGUEL, M. N. O aumento da formalização das relações de trabalho como consequência da expansão econômica. **Mercado de Trabalho – conjuntura e análise**, Rio de Janeiro: Ipea, v. 39, maio 2009 (Nota Técnica).

- DEJONG, N.; WHITEMAN, C. H. The case for trend-stationarity is stronger than we thought. **Journal of Applied Econometrics**, n. 6, p. 413-421, 1991.
- FRÜHWIRTH-SCHNATTER, S. Data augmentation and dynamic linear models. **Journal of Time Series Analysis**, n. 15, p. 183-202, 1994.
- GAMERMAN, D.; LOPES, H. **Markov chain Monte Carlo**. Chapman & Hall, 1997.
- GELMAN, A. *et al.* **Bayesian data analysis**. Chapman & Hall/CRC, 2003.
- GELMAN, A., RUBIN, D. R., Interference from iterative simulation using multiple sequences (with discussion). **Statist. Sci.**, n. 7, p. 457-511, 1992.
- HAKKIO, C.; RUSH, M. Is the budget deficit too large? **Economic Inquiry**, v. 29, n. 7, p. 429-445, 1991.
- HERNANDEZ, B. **Um modelo econométrico da conta corrente do governo no Brasil**. Rio de Janeiro: Ipea, 1998 (Texto de Discussão, n. 1951/95).
- IBGE. **Sistema de Contas Nacionais Brasil 2002-2006**. Rio de Janeiro: Coordenação de Contas Nacionais, 2008a.
- _____. **Sistema de Contas Nacionais: Brasil**. 2. ed. Rio de Janeiro, 2008b (Série Relatórios Metodológicos, v. 24).
- KIM, C. J.; NELSON, C. R. **State-space models with regime switching**. Massachusetts: The MIT Press, 1999.
- KOOP, V. G. “Objective” Bayesian unit root tests. **Journal of Applied Econometrics**, v. 7, p. 65-82, 1992.
- PORTUGAL, C.; PORTUGAL, M. Os efeitos da inflação sobre o orçamento do governo: uma análise empírica. **Estudos Econômicos**, v. 31, n. 2, p. 239-283, 2001.
- SANTOS, C. H. M.; RIBEIRO, M. B.; GOBETTI, S. W. **A evolução da carga tributária bruta brasileira no período 1995-2007: tamanho, composição e especificações econométricas agregadas**. Brasília: Ipea, 2008 (Texto para Discussão, n. 1.350).
- SCHETTINI, B. P. *et al.* **Novas evidências empíricas sobre a dinâmica trimestral do consumo agregado das famílias brasileiras no período 1995-2009**. Brasília: Ipea, 2010 (Texto para Discussão). No prelo.
- SIMS, C. Bayesian skepticism on unit root econometrics. **Journal of Economic Dynamics and Control**, v. 12, n. 2-3, p. 463-474, 1988.
- _____.; UHLIG, H. Understanding unit rooters: a helicopter tour. **Econometrica**, v. 59, n. 6, p. 1.591-1.599, 1991.
- TANZI, V. Inflation, lags in collection, and the real value of tax revenue. **International Monetary Fund Staff Papers**, Washington, v. 24, n. 1, 1977.
- WEST, M.; HARRISON, J. **Bayesian forecasting and dynamic models**. New York: Springer Verlag, 1997.

(Originais submetidos em julho de 2009. Última versão recebida em fevereiro de 2011.
Aprovada em abril de 2011.)

ANEXO A

GRÁFICO A.1

Gráfico das variáveis

Fonte: Elaboração dos autores.

ANEXO B

A METODOLOGIA DE ESTIMAÇÃO UTILIZADA NESTE TRABALHO

B.1 Modelo de regressão linear dinâmico

A forma geral do DLM pode ser definida da seguinte forma:

$$y_t = X_t \theta_t + S_t + v_t' \quad (\text{B.1})$$

$$\theta_t = \Gamma_1 \theta_{t-1} + \dots + \Gamma_p \theta_{t-p} + \omega_t' \quad (\text{B.2})$$

onde y_t é a variável endógena e X_t é um vetor $k \times 1$ de variáveis exógenas. Este modelo permite ainda a presença de um componente sazonal, S_t , que pode ser modelado segundo a representação do tipo de Fourier (WEST; HARRISON, 1997), de modo que

$$S_t = \sum_{j=1}^2 \left[a_j \sin\left(\frac{2\pi j t}{4}\right) + b_j \cos\left(\frac{2\pi j t}{4}\right) \right] = Z_t \varphi$$

onde:

$$Z_t = \left(\sin\left(\frac{2\pi t}{4}\right) \sin\left(\frac{2\pi 2t}{4}\right) \cos\left(\frac{2\pi t}{4}\right) \cos\left(\frac{2\pi 2t}{4}\right) \right)^T$$

e:

$$\varphi = (a_1 \ a_2 \ b_1 \ b_2)$$

sendo θ_t o vetor $k \times 1$ de parâmetros que segue um processo autorregressivo de ordem p , em que os coeficientes estão contidos na matriz diagonal Γ_i , $k \times k$, $i = 1, \dots, p$. Supõe-se que $\omega_t' \sim N(0_{k \times 1}, W_{k \times k})$ e $v_t' \sim N(0, V)$, admitindo-se ainda que ω_t' e v_t' não são correlacionados. Por simplicidade, admite-se também que $W = \text{diag}(w_1, \dots, w_k)$.

A ideia da aplicação da forma de Fourier para tratar sazonalidade está ligada ao fato de que qualquer padrão cíclico pode ser representado em termos da combinação

linear de funções periódicas. Outra motivação para o uso desta abordagem, além da flexibilidade, é a facilidade de se obter interpretação, já que as trajetórias em forma de ondas podem ser associadas aos variados padrões observados na prática. Mais uma forma de tratar a sazonalidade seria introduzi-la diretamente na equação de estado. Neste caso, a equação (B.2) assumiria a seguinte forma:

$$\theta_t = \Gamma_1 \theta_{t-1} + \dots + \Gamma_p \theta_{t-p} + S_t + \omega'_t \quad (\text{B.2}')$$

Por fim, a sazonalidade pode tomar uma forma dinâmica. Neste caso, seria possível ter-se um modelo no qual a equação da medida é dada da seguinte forma:

$$y_t = F_t \theta_t^* + v'_t$$

onde:

$$F = [X_t, Z_t]$$

e:

$$\theta_t^* = [\theta_t, \varphi_t]$$

Será feito agora uso da seguinte notação para encaixar as equações numa forma mais compacta, que reduz o sistema expresso pelas equações (B.1)/(B.2) numa forma de espaço-estado de ordem um:

$$\underline{\theta}_t = \begin{bmatrix} \theta_t \\ \theta_{t-1} \\ \vdots \\ \theta_{t-p+1} \end{bmatrix}_{kp \times 1}, \quad G = \begin{bmatrix} \Gamma_1 & \Gamma_2 & \dots & \Gamma_{p-1} & \Gamma_p \\ I_k & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & I_k & 0 \end{bmatrix}_{kp \times kp}, \quad \underline{\omega}_t = \begin{bmatrix} \omega'_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{kp \times 1}$$

e:

$$F'_t = \begin{bmatrix} X'_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{kp \times 1}$$

Visto isso, tem-se que:

$$y_t = F_t \underline{\theta}_t + z_t \varphi + v_t' \quad (\text{B.3})$$

$$\underline{\theta}_t = G \underline{\theta}_{t-1} + \underline{\omega}_t \quad (\text{B.4})$$

onde (B.3) representa a equação da medida ou do movimento, enquanto (B.4) define a equação de estado.

A função de verossimilhança de (θ, G, φ, V) é dada por

$$p(y/\theta, G, \varphi, V) = (2\pi)^{-T/2} V^{T/2} \exp\left(-\frac{1}{2V} \sum_{t=1}^T (y_t - F_t \underline{\theta}_t - z_t \varphi)^2\right)$$

onde:

$$y = (y_1 \dots y_T), \theta = (\underline{\theta}_1 \dots \underline{\theta}_T)$$

e:

$$\Gamma = (\Gamma_1 \dots \Gamma_p)$$

B.2 Procedimento de inferência

Nesta seção, é realizado um procedimento de inferência baseado no paradigma bayesiano. Inicialmente apresentam-se distribuições *a priori* para todos os parâmetros. Em seguida, o algoritmo MCMC é utilizado para obter amostras *a posteriori* de todos os parâmetros do modelo.

Distribuição *a priori*

Por simplicidade, distribuições *a priori* conjugadas são utilizadas para todos os parâmetros. A distribuição *a priori* para θ_t é especificada na equação (B.4) e completada pela informação inicial θ_0 , tal que $\theta_0 \sim N(m_0, C_0)$, onde m_0 e C_0 são hiperparâmetros conhecidos. As distribuições *a priori* para os parâmetros V, W, Γ e φ são as seguintes: *i*) $V \sim GI(n_V/2, n_V S_V/2)$; *ii*) $w_j \sim GI(n_W/2, n_W S_W/2), j = 1, \dots, k$; *iii*) $\Gamma_j \sim N(m_\Gamma, S_\Gamma), j = 1, \dots, p$; e *iv*) $\varphi \sim N(m_\varphi, S_\varphi)$, onde $n_V, S_V, n_W, S_W, m_\Gamma, S_\Gamma, m_\varphi$ e S_φ são hiperparâmetros conhecidos.

Os valores dos hiperparâmetros utilizados foram: $m_0 = 0$, $C_0 = 100$, $n_V = 2$, $n_V S_V = 0.1$, $n_W = 2$, $n_W S_W = 0.1$, $m_\Gamma = 0$, $S_\Gamma = 100$, $m_\phi = 0$, $S_\phi = 100$.

Inferência *a posteriori*

A distribuição *a posteriori* conjunta de $(\theta, \Gamma, \phi, V, W)$ é dada por:

$$p(\theta, G, \phi, V, W | y) \propto \prod_{t=1}^T p(y_t | \theta_t, \phi, V) \\ \prod_{t=1}^T p(\theta_t | \theta_{t-1}, W, G) p(\theta_0 | m_0, C_0) P(\phi) p(V) p(W) p(\Gamma)$$

a qual é analiticamente intratável e, por conseguinte, a inferência *a posteriori* exata é feita utilizando-se o esquema MCMC. Neste caso, o vetor da variável de estado é amostrado conjuntamente, utilizando-se o algoritmo Forward Filtering Backward Sampling (FFBS) de Carter e Kohn (1994) e Frühwirth-Schnatter (1994). Para o resto de parâmetros Γ , ϕ , V e W , todas as distribuições condicionais completas são distribuições normais ou distribuições gama inversa. Todas as distribuições condicionais completas estão listadas a seguir:

- A distribuição condicional de V é:

$$(V | \theta, W, \phi, \Gamma) \sim GI(n_V^* / 2, n_V^* S_V^* / 2)$$

onde:

$$n_V^* = n_V + T$$

e:

$$n_V^* S_V^* = n_V S_V + \sum_{t=1}^T (y_t - F_t \theta_t - z_t \phi)^2$$

- A distribuição condicional de $W = \text{diag}(w_1, \dots, w_k)$ é:

$$(w_j | \theta, V, \phi, \Gamma) \sim GI(n_W^* / 2, n_W^* S_W^* / 2) \text{ para } j = 1, \dots, k$$

onde:

$$n_W^* = n_W + T - 1$$

e:

$$n_W^* S_W^* = n_W S_W + \sum_{t=2}^T (\underline{\theta}_t - G \underline{\theta}_{t-1})^T (\underline{\theta}_t - G \underline{\theta}_{t-1})$$

- A distribuição condicional de $\Gamma = (\Gamma_1 \dots \Gamma_p)$ é:

$$(\Gamma_j | \theta, V, W, \varphi) \sim N(m_\Gamma^*, S_\Gamma^*) \text{ para } j = 1, \dots, k$$

onde:

$$S_\Gamma^* = (S_\Gamma^{-1} I_{p \times p} + w_j^{-1} H_j^T H_j)^{-1}$$

$$m_\Gamma^* = S_\Gamma^{-1} (S_\Gamma^{-1} I_{p \times p} \mathbf{1}_p + w_j^{-1} B_j^T H_j)$$

para:

$$H_j = \begin{pmatrix} \theta_{p,j} & \theta_{p-1,j} & \cdots & \theta_{1,j} \\ \theta_{p+1,j} & \theta_{p,j} & \cdots & \theta_{2,j} \\ \vdots & \vdots & \vdots & \vdots \\ \theta_{T-1,j} & \theta_{T-2,j} & \cdots & \theta_{T-p,j} \end{pmatrix}$$

e:

$$B_j = \begin{pmatrix} \theta_{p+1,j} \\ \theta_{p+2,j} \\ \vdots \\ \theta_{T,j} \end{pmatrix}$$

- A distribuição condicional de φ é:

$$(\varphi | \theta, V, W, \Gamma) \sim N(m_\varphi^*, S_\varphi^*)$$

onde:

$$S_\varphi^* = (S_\varphi^{-1} I_{4 \times 4} + V^{-1} Z^T Z)^{-1}$$

e:

$$m_\varphi^* = S_\varphi^{-1} (S_\varphi^{-1} I_{4 \times 4} 1_4 + V^{-1} Z^T (y - F\theta))$$

Os estados $\theta_1, \dots, \theta_T$ são amostrados conjuntamente, utilizando-se o algoritmo FFBS, condicional sobre Γ , φ , V e W . A distribuição condicional completa conjunta segue a decomposição retrospectiva $p(\theta | y) = p(\theta_T | D_T) \prod_{t=1}^{T-1} p(\theta_t | \theta_{t+1}, D_t)$, onde $D_T = \{y_1, y_2, \dots, y_T\}$, $t = 1, \dots, T$ e D_0 representam a informação inicial. Iniciando com $\theta_0 \sim N(m_0, C_0)$, através do filtro de Kalman pode ser mostrado que $\theta_t | D_t \sim N(m_t, C_t)$, onde $m_t = a_{t-1} + A(y_t - f_t - z_t \varphi)$, $C_t = R_t - A_t Q_t A_t'$, $a_t = G m_{t-1}$, $R_t = G C_{t-1} G' + W$, $f_t = F_t' a_t + z_t \varphi$, $Q_t = F_t' R_t F_t + V$ e $A_t = R_t F_t Q_t^{-1}$ para $t = 1, \dots, T$; θ_T é amostrado de $p(\theta_T | D_T)$ (passo denominado *forward filtering*). Para $t = T-1, T-2, \dots, 2, 1$, θ_t é amostrado de $p(\theta_t | \theta_{t+1}, D_t) = N(\tilde{a}_t, \tilde{C}_t)$, onde $\tilde{a}_t = m_t + B_t (f_{t+1} - a_{t+1})$, $\tilde{C}_t = C_t - B_t R_{t+1} B_t'$ e $B_t = C_t G' R_{t+1}^{-1}$ (passo denominado *backward sampling*).

Previsão de modelo de regressão linear dinâmico

Fazer previsão nesta classe de modelos é teoricamente direto. Frequentemente, surge o interesse de aprender a densidade preditiva h passos à frente, $p(y_{T+h} | y)$, ou seja,

$$p(y_{T+h} | y) = \int p(y_{T+h} | \underline{\theta}_{T+h}, \varphi, V) p(\underline{\theta}_{T+h} | \underline{\theta}_T, W, \Gamma) p(\underline{\theta}_T, W, \Gamma, V | y) d\underline{\theta}_{T+h} d\underline{\theta}_T d\varphi dV dW d\Gamma$$

onde:

$$(y_{T+b} | \underline{\theta}_{T+b}, \varphi, V) \sim N(F_{T+b}, V), (\underline{\theta}_{T+b} | \underline{\theta}_T, W, \Gamma) \sim N(\mu_b, V_b), \mu_b = G^b \underline{\theta}_T$$

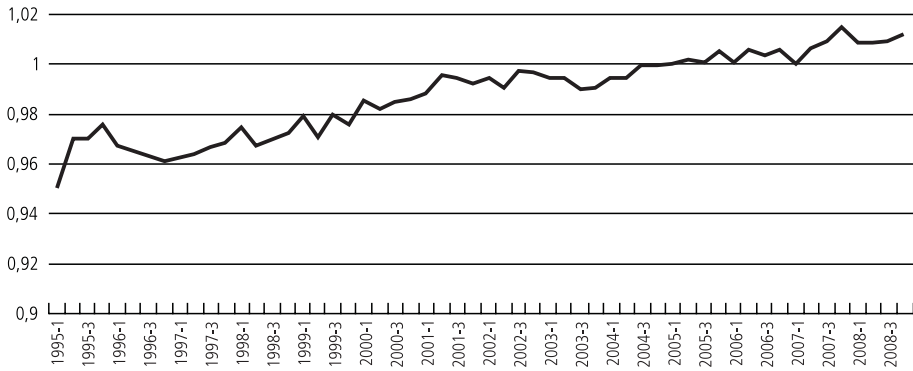
e:

$$V_b = \sum_{j=1}^b G^{j-1} W (G^{j-1}) \quad \text{para } b > 0$$

Portanto, se $\{(\underline{\theta}_T^{(1)}, \varphi^{(1)}, V^{(1)}, W^{(1)}, \Gamma^{(1)}), \dots, (\underline{\theta}_T^{(L)}, \varphi^{(L)}, V^{(L)}, W^{(L)}, \Gamma^{(L)})$ é uma amostra de $p(\underline{\theta}_T, W, \Gamma, V | y)$, é fácil amostrar $\underline{\theta}_{T+b}^{(j)}$ de $p(\underline{\theta}_{T+b} | \underline{\theta}_T^{(j)}, W^{(j)}, \Gamma^{(j)})$, para todo $j = 1, \dots, L$, tal que $p(y_{T+b} | y) = \frac{1}{L} \sum_{j=1}^L p(y_{T+b} | \underline{\theta}_{T+b}^{(j)}, \varphi^{(j)}, V^{(j)})$ é a aproximação via Monte Carlo de $p(y_{T+b} | y)$. Analogamente, se $y_{T+b}^{(j)}$ é amostrado de $p(y_{T+b} | \underline{\theta}_{T+b}^{(j)}, \varphi^{(j)}, V^{(j)})$, para $j = 1, \dots, L$, então $\{y_{T+b}^{(1)}, \dots, y_{T+b}^{(L)}\}$ representa uma amostra de $p(y_{T+b} | y)$.

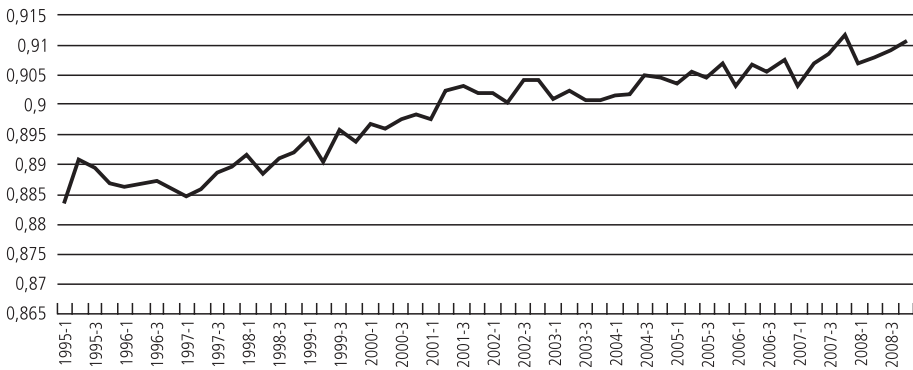
ANEXO C

GRÁFICO C.1
Evolução da mediana *a posteriori* de β_1 (INDPIB)



Fonte: Elaboração dos autores.

GRÁFICO C.2
Evolução da mediana *a posteriori* de β_1 (PIB)



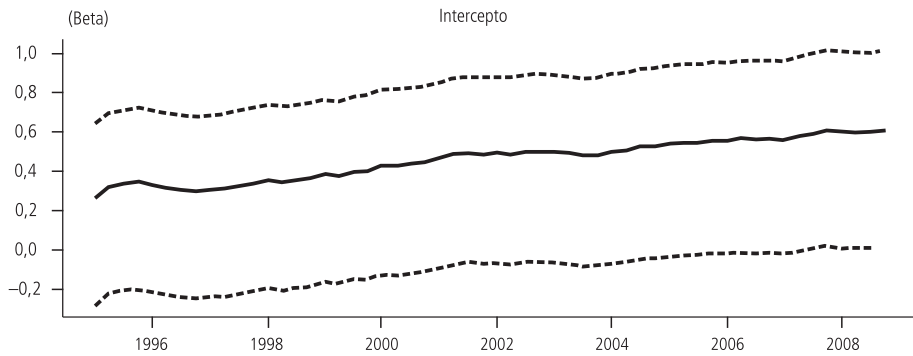
Fonte: Elaboração dos autores.

ANEXO D

RESULTADOS DOS MODELOS 3 a 5

GRÁFICO D.1

Modelo 3: INDPIB fixo

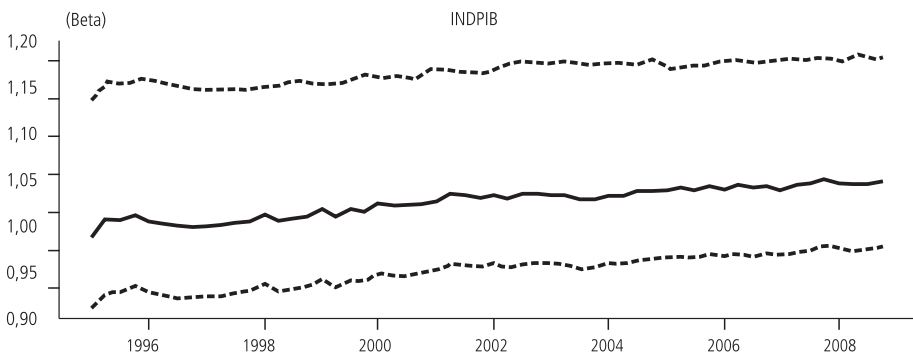


θ	$E(\theta)$	IC 95%	
Beta 1	0,943521	0,8612248	1,060735

Fonte: Elaboração dos autores.

GRÁFICO D.2

Modelo 4: constante fixa



θ	$E(\theta)$	IC 95%	
Intercepto	0,123943	-0,71426	0,529692

Fonte: Elaboração dos autores.

Modelo 5: constante fixa e INDPIB fixo

θ	$E(\theta)$	IC 95%	
Intercepto	-3.357206	-3.928881	-2.793624
Beta 1	1.746078	1.627032	1.867461

Fonte: Elaboração dos autores.

ANEXO E

TABELA E.1

Dados trimestrais referentes ao período janeiro de 1995-dezembro de 2009

	CTBB	PIB	INDPIB	Inflação	DLSP
1995-1	10.72	11.96	4.56	4.92	6.42
1995-2	10.78	12.05	4.60	6.28	6.44
1995-3	10.77	12.10	4.63	5.67	6.50
1995-4	10.86	12.19	4.63	3.88	6.57
1996-1	10.82	12.04	4.56	3.64	6.58
1996-2	10.75	12.09	4.61	3.03	6.66
1996-3	10.79	12.16	4.68	2.64	6.70
1996-4	10.78	12.12	4.64	0.92	6.73
1997-1	10.83	12.07	4.59	2.10	6.74
1997-2	10.78	12.14	4.65	1.83	6.72
1997-3	10.82	12.17	4.69	0.72	6.75
1997-4	10.85	12.16	4.67	0.52	6.80
1998-1	10.92	12.08	4.60	1.47	6.85
1998-2	10.82	12.15	4.67	0.96	6.91
1998-3	10.85	12.17	4.69	-0.35	6.95
1998-4	10.87	12.14	4.66	-0.27	7.01
1999-1	10.96	12.09	4.61	1.94	7.13
1999-2	10.83	12.14	4.67	1.90	7.15
1999-3	10.91	12.16	4.68	1.79	7.16
1999-4	10.91	12.16	4.68	2.41	7.12
2000-1	11.04	12.13	4.65	1.49	7.13
2000-2	10.94	12.18	4.70	0.69	7.13
2000-3	10.96	12.20	4.72	2.70	7.10
2000-4	11.01	12.20	4.72	1.13	7.11
2001-1	11.07	12.17	4.68	1.50	7.14
2001-2	11.05	12.21	4.73	1.43	7.16
2001-3	11.03	12.20	4.73	2.36	7.21

(continua)

(continuação)

	CTBB	PIB	INDPIB	Inflação	DLSP
2001-4	11.05	12.20	4.72	1.93	7.20
2002-1	11.12	12.17	4.69	1.63	7.21
2002-2	11.03	12.22	4.75	1.59	7.26
2002-3	11.08	12.24	4.77	2.20	7.35
2002-4	11.12	12.24	4.77	4.69	7.23
2003-1	11.14	12.19	4.71	6.03	7.20
2003-2	11.08	12.23	4.76	2.65	7.17
2003-3	11.05	12.25	4.78	0.79	7.20
2003-4	11.09	12.25	4.78	1.32	7.21
2004-1	11.20	12.24	4.77	1.78	7.19
2004-2	11.14	12.31	4.83	1.46	7.18
2004-3	11.16	12.30	4.82	2.11	7.15
2004-4	11.19	12.30	4.82	1.63	7.15
2005-1	11.26	12.27	4.80	1.97	7.14
2005-2	11.23	12.33	4.86	1.79	7.14
2005-3	11.18	12.33	4.85	0.63	7.16
2005-4	11.26	12.33	4.85	1.52	7.19
2006-1	11.30	12.32	4.84	1.43	7.21
2006-2	11.27	12.35	4.88	0.63	7.20
2006-3	11.25	12.38	4.90	0.19	7.21
2006-4	11.32	12.38	4.90	0.85	7.22
2007-1	11.34	12.37	4.89	1.28	7.23
2007-2	11.32	12.42	4.94	0.92	7.24
2007-3	11.33	12.43	4.95	0.89	7.23
2007-4	11.43	12.44	4.96	1.08	7.22
2008-1	11.45	12.43	4.94	1.64	7.19
2008-2	11.39	12.48	5.00	1.80	7.17
2008-3	11.40	12.50	5.02	1.55	7.13
2008-4	11.41	12.45	4.97	1.05	7.10
2009-1	11.43	12.41	4.93	1.22	7.13
2009-2	11.35	12.46	4.98	1.23	7.20
2009-3	11.35	12.49	5.00	0.81	7.25
2009-4	11.44	12.49	5.00	0.89	7.26

Fonte: Elaboração dos autores.

Ipea – Instituto de Pesquisa Econômica Aplicada

Editorial

Coordenação

Claudio Passos de Oliveira

Supervisão

Andrea Bossle de Abreu

Revisão

Eliezer Moreira

Elisabete de Carvalho Soares

Fabiana da Silva Matos

Gilson Baptista Soares

Lucia Duarte Moreira

Luciana Nogueira Duarte

Míriam Nunes da Fonseca

Editoração

Roberto das Chagas Campos

Aeromilson Mesquita

Camila Guimarães Simas

Carlos Henrique Santos Vianna

Maria Hosana Carneiro da Cunha

Paula Mascarenhas Rodrigues de Almeida (estagiária)

Brasília

SBS – Quadra 1 – Bloco J – Ed. BNDES,

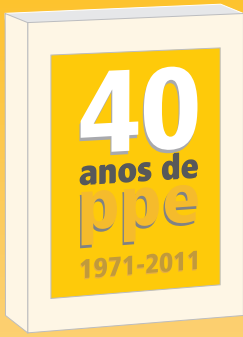
Térreo – 70076-900 – Brasília – DF

Fone: (61) 3315-5336

Correio eletrônico: livraria@ipea.gov.br

NOTA AOS COLABORADORES DE PESQUISA E PLANEJAMENTO ECONÔMICO

1. A revista só analisa, com vistas a eventual publicação, artigos com conteúdo inédito, tanto no país quanto no exterior. Além disso, o seu tema deve se inserir em uma das áreas da ciência econômica, contribuindo de modo significativo ao avanço do conhecimento científico nessa área.
2. Resenhas de livros recentemente publicados poderão ser consideradas para publicação, mas resenhas temáticas e os textos essencialmente descritivos não serão, de um modo geral, aceitos.
3. As contribuições não serão remuneradas, e a submissão de um artigo à revista implica a transferência dos direitos autorais ao Ipea, caso ele venha a ser publicado.
4. Em geral, os artigos submetidos à revista devem ser escritos em português. Em casos excepcionais, poderão ser recebidos textos em língua inglesa para análise, mas se ele vier a ser aceito para publicação, o autor deverá se responsabilizar por sua tradução.
5. Só serão publicados artigos em português, mas sua versão em inglês poderá ser disponibilizada no sítio da revista na internet. Os anexos muito longos ou complexos para serem publicados, bem como as bases de dados necessárias para reproduzir os resultados empíricos do trabalho, serão também oferecidos aos leitores em versão virtual.
6. Caso o trabalho seja aceito para publicação, cada autor receberá 3 (três) exemplares do número da revista correspondente.
7. Para submeter um trabalho à revista, o autor deve acessar a página de Pesquisa e Planejamento Econômico na internet, em <http://ppe.ipea.gov.br/index.php/ppe/index>, e seguir os procedimentos ali descritos.
8. Os artigos recebidos que estejam de acordo com as instruções acima serão avaliados pelo Corpo Editorial com o auxílio de dois pareceristas escolhidos pelo Editor. O trabalho dos pareceristas é feito observando-se o método duplamente cego: o autor não saberá quem são os pareceristas, nem estes quem é o autor. Dessa análise poderá resultar a aceitação do artigo, condicionada, ou não, à realização de alterações; sua rejeição, com ou sem a recomendação de nova submissão após modificações; ou a rejeição definitiva. No caso de uma segunda submissão, o artigo será novamente avaliado por pareceristas, podendo vir a ser enquadrado em qualquer das situações acima. A rotina de análise se repete até que uma decisão final de rejeição ou aceitação seja alcançada. O processamento do artigo é conduzido pelo Editor, a quem cabe também a comunicação com os autores.
9. A decisão final quanto à publicação dos artigos cabe ao Corpo Editorial, que se reúne ordinariamente para decidir a composição de cada um dos números da revista, por recomendação do Editor. A aprovação do artigo para publicação só então é comunicada aos autores dos artigos respectivos, por escrito.



ISSN 0100-0551



Apoio editorial



Ipea - Instituto de
Pesquisa Econômica
Aplicada

