

Anais do II Encontro de Economia Aplicada

22 a 23 de junho de 2016



**Programa de Pós-graduação em Economia
Faculdade de Economia
Universidade Federal de Juiz de Fora**

Comissão Organizadora

Admir Antônio Bertarelli Júnior, Universidade Federal de Juiz de Fora
Carlos Henrique Leite Corseuil, Instituto de Pesquisas Econômicas Aplicadas
Cláudio Roberto Fóffano Vasconcelos, Universidade Federal de Juiz de Fora
Edson Paulo Domingues, Cedeplar – Universidade Federal de Minas Gerais
Eduardo Amaral Haddad, FEA – Universidade de São Paulo
Eduardo Gonçalves, Universidade Federal de Juiz de Fora
Eduardo Simões de Almeida, Universidade Federal de Juiz de Fora
Enlinson Henrique Carvalho de Mattos, Fundação Getúlio Vargas – São Paulo
Fernando Salgueiro Perobelli, Universidade Federal de Juiz de Fora
Flávia Lúcia Chein Feres, Universidade Federal de Juiz de Fora
José Simão Filho, Universidade Federal de Juiz de Fora
Marcel de Toledo Vieira, Universidade Federal de Juiz de Fora
Marcelo Aarestrup Arbex, Universidade de Windsor – Canadá
Ricardo da Silva Freguglia, Universidade Federal de Juiz de Fora (Coordenador)
Sidney Martins Caetano, Universidade Federal de Juiz de Fora
Wilson Luiz Rotatori Corrêa, Universidade Federal de Juiz de Fora

Comissão Científica

Área 1: Economia Regional (Regional Economics)

Edson Paulo Domingues, Cedeplar – Universidade Federal de Minas Gerais
Eduardo Amaral Haddad, FEA – Universidade de São Paulo
Fernando Salgueiro Perobelli, Universidade Federal de Juiz de Fora

Área 2: Economia Social e do Trabalho (Labor and Social Economics)

Cristine Campos de Xavier Pinto, Fundação Getúlio Vargas – São Paulo
Kenya Valeria Micaela de Souza Noronha, Cedeplar – Universidade Federal de Minas Gerais
Flávia Lúcia Chein Feres, Universidade Federal de Juiz de Fora

Área 3: Microeconomia Aplicada: Finanças e Organização Industrial (Applied Microeconomics: Finance and Industrial Organization)

Claudia Emiko Yoshinaga, Fundação Getúlio Vargas – São Paulo
Fernanda Finotti Cordeiro Perobelli, Universidade Federal de Juiz de Fora
Francisco Henrique Figueiredo de Castro Junior, Universidade de São Paulo

Área 4: Macroeconomia Aplicada (Applied Macroeconomics)

Gabriel Caldas Montes, Universidade Federal Fluminense
José Ronaldo Castro Souza Junior, Instituto de Pesquisas Econômicas Aplicadas
Wilson Luiz Rotatori Corrêa, Universidade Federal de Juiz de Fora

Sessão 1: Economia Regional (Regional Economics)

Pobreza Rural e Dualidade Produtiva no Nordeste Brasileiro

Felipe Pinto da Silva (MAER/UFC), Jair Andrade Araújo (MAER/UFC), Edward Martins Costa (MAER/UFC) e José Eustáquio Ribeiro Vieira Filho (Dirur/Ipea)

Wages and Net Fiscal Benefit in a Federation: Evaluating Intergovernmental Grants in Brazilian Municipalities

Ricardo Batista Politi (UFABC) e Enlinson Mattos (EESP – FGV)

Sessão 2: Macroeconomia Aplicada (Applied Macroeconomics)

Transparência fiscal, comprometimento do governo e credibilidade fiscal: Uma análise considerando os efeitos da contabilidade criativa no Brasil

Tatiana Acar (UFF)

Prevendo a inflação no Brasil com grande conjunto de dados: uma aplicação do modelo de fatores comuns

Cristiano Santos (UFAL), Roberto Tatiwa Ferreira (CAEN/UFC), Ivan Castelar (CAEN/UFC)

Comunicação Macropprudencial: Evidências do efeito da comunicação do banco central sobre a estabilidade do sistema financeiro

Claudio Oliveira de Moraes (Bacen e UFF) e Rodolfo Nicolay (UCAM)

Session 3: Microeconomia Aplicada: Finanças e Organização Industrial (Applied Microeconomics: Finance and Industrial Organization)

Internet diffusion: killing time prevents crime

Ilaria Masiero (PhD Candidate at FGV-EESP)

Mobilidade dos Fatores Primários e seus Efeitos sobre o Crescimento Econômico e o Bem-Estar Gerados pelo Crédito Rural nas Regiões Brasileiras

Talita Priscila Pinto (UFV), Eryl Cardoso Teixeira (UFV), Ângelo Costa Gurgelc (FVG/SP)

The Innovative Aspects Of Bitcoin, Market Microstructure and Returns Volatility: an Approach Using Mgarch

Bruno Ferreira Frascaroli (PPGE/UFPB) e Thiago Carvalho Pinto (PPGE/UFPB)

Sessão 4: Economia Social e do Trabalho (Labor and Social Economics)

Impactos do Programa Bolsa Familia no Mercado de Trabalho e na Renda dos Trabalhadores Rurais

Edward Martins Costa (MAER/UFC), Rayssa Alexandre Costa (ESALQ/USP), Francisca Zilania Mariano (CAEN/UFC) e Daniella Medeiros Cavalcante (PIMES/UFPE)

Heat Shocks, Child Endowments, and Parental investments

Bladimir Carrillo (UFV), João Lima (UFV) e Juan C. Trujillo (University of York Heslington)

New Evidence of the Effect of Body Weight on Labor Market Outcomes in Brazil

Carlos A. Charris (UFV), Bladimir Carrillo (UFV), Pedro Vertino (UFV) e Alexandre B. Coelho (UFV)

Secondary vocational education and earnings in Brazil

Marina Aguas (IBGE) e Danielle Carusi Machado (UFF)

POBREZA RURAL E DUALIDADE PRODUTIVA NO NORDESTE BRASILEIRO

Felipe Pinto da Silva¹

Jair Andrade Araújo²

Edward Martins Costa³

José Eustáquio Ribeiro Vieira Filho⁴

RESUMO

Este estudo busca analisar as diferenças tecnológicas da produção agropecuária entre os municípios situados no semiárido e no não-semiárido do Nordeste brasileiro. Utiliza-se o modelo de meta-fronteira tecnológica, proposta por Battese *et al.* (2004) e O'Donnell *et al.* (2008), para responder até que ponto essas diferenças regionais influenciam na eficiência da produção e qual o seu impacto nos municípios. Constatou-se que a hipótese de que os municípios pertencentes ao semiárido e ao não-semiárido deparam-se com diferentes oportunidades de produção. Os resultados indicaram que, quando comparadas as duas regiões, o não-semiárido possui a maior eficiência técnica média com referência a meta-fronteira.

PALAVRAS CHAVES: Agricultura, tecnologia, semiárido nordestino, e meta-fronteira de produção.

Classificação JEL: D24, Q10.

RURAL POVERTY AND PRODUCTIVE DUALITY IN NORTHEASTERN BRAZIL

ABSTRACT

This study aims analyze the technological differences in agricultural production between the municipalities in the semiarid and non-semiarid Northeast Brazil. It uses the technological meta-frontier model proposed by Battese *et al.* (2004), and O'Donnell *et al.* (2008), for responding to what extent these regional differences influence the efficiency of production and what its impact on municipalities. It was observed that the hypothesis that the municipalities belonging to the semiarid and non-semiarid face with different production opportunities. The results indicated that compared the two regions, the non-semiarid region has the highest average technical efficiency with reference to meta-frontier.

KEY WORDS: Agriculture, technology, semi-arid northeast, and production meta-frontier.

JEL classification: D24, L23, Q10.

¹ Mestrando em Economia Rural (MAER/UFC). E-mail: dasilva.felipe@outlook.com

² Doutor em Economia (CAEN/UFC) e Professor do Curso de Mestrado em Economia Rural (MAER/UFC). E-mail: jaraujoce@gmail.com

³ Doutor em Economia (PIMES/UFPE) e Professor do Curso de Mestrado em Economia Rural (MAER/UFC). E-mail: edwardcosta@ufc.br

⁴ Técnico de Planejamento e Pesquisa da Diretoria de Estudos Regionais, Urbanos e Ambientais (Dirur) do Instituto de Pesquisa Econômica Aplicada (Ipea). E-mail: jose.vieira@ipea.gov.br.

1. INTRODUÇÃO

O Brasil passou por um processo intenso de modernização de seu setor agrícola a partir da década de 1960. Gasques *et al.* (2012) apresentam evidências do crescimento elevado da produtividade agrícola nacional, com taxas superiores aos principais países produtores de grãos e carnes, tais como Estados Unidos, União Europeia, China, Índia e Argentina. De forma complementar, Fugile e Wang (2012) constataam que a agricultura brasileira e a chinesa apresentaram, nos últimos 20 anos, o maior e mais rápido crescimento da produtividade total dos fatores (PTF) em âmbito mundial. Além disso, essa tendência se intensificou durante o período de 2000 a 2012 (GASQUES, 2014).

Nota-se que esse processo se caracterizou de maneira heterogênea entre as regiões brasileiras. O Nordeste merece atenção nessa trajetória por não ter acompanhado a dinâmica observada no país, tornando-se a região brasileira com menor dinamismo no setor agrícola (VICENTE, 2004; VIEIRA FILHO, 2013). No Brasil, é fácil perceber a convivência dual do moderno *versus* o atrasado. Enquanto que nas regiões Centro-Oeste e Sul encontra-se um setor agrícola moderno, com alto teor tecnológico e elevada produtividade, no eixo Norte e Nordeste, há um contraste de uma agricultura em que prevalece produtores que utilizam técnicas atrasadas e com baixa produtividade. A situação da pobreza rural se agrava no Nordeste brasileiro (VIEIRA FILHO, 2013).

O Nordeste é uma das áreas mais vulneráveis do Brasil às variabilidades climáticas, que se associam às chuvas irregulares, aos déficits hídricos e à baixa resiliência das populações que se dedicam à agricultura, principalmente nos assentamentos da região semiárida (SIMÕES *et al.*, 2010). Estas características fazem da agropecuária um empreendimento cercado de riscos, especialmente no semiárido. Ao contrário da dinâmica da moderna produção, o setor agropecuário constituído no semiárido nordestino tem acesso restrito aos mercados, o que dificulta o desenvolvimento regional. Vale observar que, mesmo dentro do Nordeste, o setor agrícola não se comporta de maneira homogênea. Agricultores situados no semiárido, por hipótese, deparam-se com condições geográficas, culturais, históricas e institucionais distintas, que não devem ser negligenciadas das análises das políticas públicas de desenvolvimento regional. É necessário promover as instituições de forma que não haja o aprofundamento da discrepância produtiva entre a região dinâmica e atrasada, como bem demonstram os estudos de Hayami (1969) e Hayami e Ruttan (1970; 1971).

Diante dessas condicionalidades, é fato que os agricultores situados no nordeste semiárido se deparam com diferentes oportunidades de produção. Assim, podem-se levantar alguns problemas: i) essas diferenças influenciam na eficiência da produção; e ii) se positiva a resposta ao primeiro questionamento, qual a magnitude dos impactos produtivos em pertencer ou não ao semiárido nordestino. Espera-se que esta problemática seja respondida ao longo do estudo.

Tendo exposta as questões centrais, como objetivo geral de investigação, busca-se identificar as diferenças tecnológicas da produção agropecuária entre os municípios situados no semiárido e nas demais regiões do Nordeste brasileiro. Para tanto, propõe-se a utilização do método utilizado por Battese *et al.* (2004) e O'Donnell *et al.* (2008) para estimação da meta-fronteira de produção. Este quadro teórico é bastante utilizado para a comparação de níveis de eficiência e tecnologias de produção entre regiões distintas.

Serão observadas em que proporção as diferenças regionais próprias do Nordeste, no que tange a qualidade do solo, diferença climática, infraestrutura econômica, dentre outras variáveis, impactam na eficiência das firmas situadas nos diferentes grupos regionais. Particularmente, a análise se concentra no estudo do Nordeste, incluindo o semiárido e as demais regiões, a partir da cartografia delimitada pela Superintendência do Desenvolvimento do Nordeste (SUDENE). Como objetivo específico, procura-se mensurar e analisar a eficiência

técnica (ET) e a razão de meta-fronteira (MRT) dos municípios inseridos no semiárido e não-semiárido nordestino.

A principal contribuição deste estudo consiste em uma melhor caracterização dos municípios situados no semiárido e nas demais regiões do Nordeste brasileiro, atentando-se para a questão produtiva e tecnológica. Não há dúvidas de que estudos técnicos que possam identificar os reais problemas da região são importantes para a formulação de políticas públicas de desenvolvimento regional, especialmente no fomento do semiárido.

Para a análise, serão apresentadas 5 seções, incluindo esta breve introdução. Na seção 2, faz-se uma breve caracterização do semiárido, além de copilar estudos empíricos relacionados à eficiência produtiva da agropecuária da região. A seção 3, apresenta a metodologia e o modelo econométrico, bem como a fonte dos dados e a amostra dos municípios usados na estimação. Na seção 4, há a discussão dos resultados. Por fim, seguem as considerações finais.

2. CARACTERIZAÇÃO DO SEMIÁRIDO DO NORDESTE BRASILEIRO

2.1. O semiárido brasileiro como delimitação empírica

O Nordeste brasileiro corresponde a uma área total de 1,57 milhões de km², perfazendo 18% do território nacional. Trata-se da segunda região de planejamento mais populosa do país, constituindo 28% da população total do Brasil e concentrando cerca de 40% da pobreza. Em 2010, essa disparidade foi ilustrada ao comparar o índice de desenvolvimento humano (IDH⁵) de 0,57 da região Nordeste com o indicador da região Sudeste, que constitui o número de 0,78, como aponta Barbieri *et al* (2010).

Nessa região, a variabilidade climática é alta, e longos períodos de secas estão associados historicamente a fome e mortalidade (DAVIS, 2002). Parcela significativa da população consiste de agricultores familiares tradicionais de subsistência e com baixo teor tecnológico. Campos e Kahn (1992) apontam que, em longos períodos de seca, as perdas no total da produção dos pequenos agricultores são de até 80%. Nasuti *et al.* (2013), ao pesquisar cerca de 1140 produtores rurais, dos estados da Bahia, do Ceará, do Pernambuco, do Rio grande do Norte e do Piauí, no período de 2011 a 2013, chegaram a resultados semelhantes, sendo que cerca de 70% dos produtores afirmavam obter prejuízos devidos à variabilidade climática.

Como já destacado, a região Nordeste é heterogênea, tanto economicamente, quanto do ponto de vista da distribuição dos recursos naturais. A região semiárida, ocupando 60% de toda região Nordeste, estende-se a uma área de 982.563,3 km², sendo que 89,5% pertence ao Nordeste brasileiro e o restante ao norte do Estado de Minas Gerais. Em termos populacionais, o semiárido nordestino compreende aproximados 12% da população do país. Popularmente conhecida como o “polígono das secas”, é a região mais pobre do Brasil, sendo sujeita a severas secas periódicas (DA MATA; RESENDE, 2015), tornando-a a região semiárida dos trópicos mais populosa do mundo (MORENGO, 2008).

Ainda em 2010, a região semiárida apresenta um nível de analfabetismo três vezes maior que a média nacional (ROCHA e RESENDE, 2015). Conforme a Tabela 1, ao comparar os indicadores demográficos, para o ano de 2010, a média da renda domiciliar *per capita* do semiárido nordestino foi 2,4 vezes menor em relação à média nacional, e apenas 70% da renda média do Nordeste.

⁵Índice sintético calculado pela média simples de três sub índices, referentes às dimensões de Longevidade (IDHM-L), Educação (IDHM-E) e Renda (IDHM-R). Para obter informações metodológicas, consulte o Atlas do Desenvolvimento Humano no Brasil - 1998. O IDH é uma medida comparativa para classificar diferentes regiões pelo grau de desenvolvimento. Quanto mais próximo de 1, mais elevado é o grau de desenvolvimento. Ao contrário, quanto menor for o indicador, mais frágil é o desenvolvimento humano da região.

Ao longo do período, de 2000 a 2010, percebe-se que o semiárido foi a região analisada com maior variação percentual de renda domiciliar *per capita*, o que resultou, consequentemente numa melhora da distribuição de renda comparada às demais regiões. Deve-se ter cuidado ao interpretar o índice de Gini⁶, pois, embora a região semiárida tenha aumentada a renda ao longo do período, os valores absolutos se mostram mais baixos do que nas demais regiões, ou seja, a maior igualdade de renda se dá em uma situação de *socialização da pobreza*, ainda que com uma melhora de renda.

Tabela 1 - Média da renda domiciliar *per capita* real e índice de Gini para o Brasil, Nordeste e Semiárido Nordestino (2000 e 2010)

Unidades	Renda (Em Reais)			Índice de Gini		
	2000	2010	Variação (%)	2000	2010	Variação (%)
Semiárido Nordestino	246,86	399,48	61,82	0,58	0,55	-5,17
Nordeste	365,89	570,10	55,81	0,64	0,62	-3,13
Brasil	701,71	949,99	35,38	0,63	0,60	-4,29

Fonte: Elaboração com base em Irff (2015).

Em relação às questões sociais, Rocha e Resende (2015) mostram que, no semiárido nordestino, no período de 1996 a 2008, aproximadamente 80% das crianças estavam a baixo da linha da pobreza e a mortalidade infantil era de 31 óbitos por 1000 nascimentos. Em contrapartida, a média brasileira, se encontrava, respectivamente, nos patamares de 25% e aproximadamente 15 óbitos por 1000 nascimentos. Todos esses indicativos tornam a região semiárida brasileira o maior bolsão de pobreza do país.

A distribuição populacional da região não acompanhou o processo de urbanização que ocorreu no Brasil no século passado (ALVES e ROCHA, 2010). De acordo com Alves e Rocha (2010), o Nordeste é a região com maior potencial de êxodo rural nos próximos anos e, por isso, é necessário definir políticas que possam fomentar o espaço rural e o urbano, de modo a não se repetir os erros do passado.

Os municípios que compõem o semiárido são pequenos, com uma população média de 12 mil habitantes, tendo sua economia fortemente baseada em formas extensivas de agricultura de subsistência com baixa produtividade e pouca resiliência às flutuações climáticas (AB'SABER, 1999; BARBIERI, *et al.*, 2010). A partir dos dados apresentados pelo Censo Agropecuário de 2006 (IBGE, 2006), a agricultura familiar é o segmento predominante no contexto produtivo da agropecuária da região.

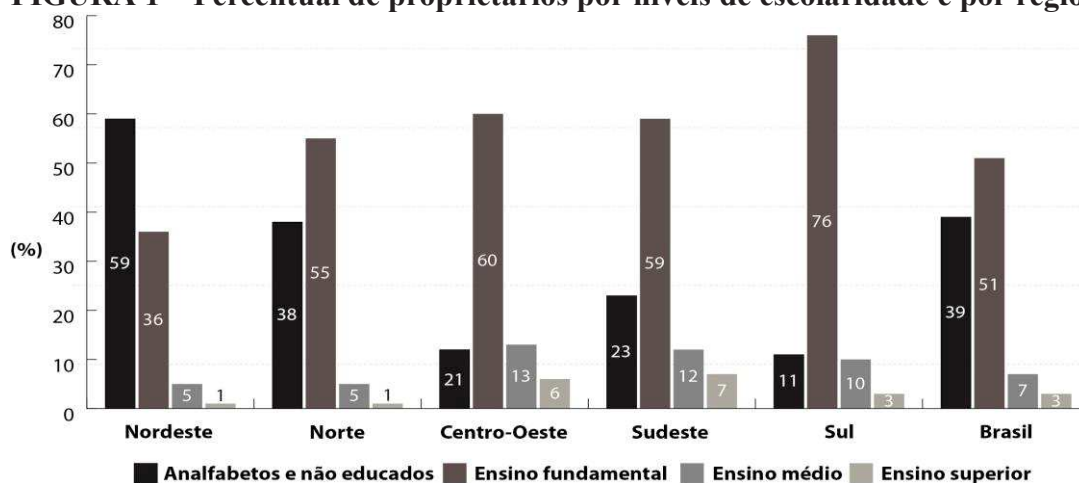
De acordo com Vieira Filho (2013), dos 5,1 milhões de estabelecimentos agropecuários no Brasil, cerca de 84% eram estabelecimentos de agricultura familiar. No comparativo regional, dos estabelecimentos de agricultura familiar, quase a metade (50%) estava localizada no Nordeste. No que se refere à extrema pobreza, esse percentual ainda é muito mais elevado na região nordestina, cerca de 61%.

No que tange aos níveis de escolaridade (Figura 1), no Brasil, 90% dos proprietários dos estabelecimentos não possuíam o ensino fundamental (analfabetos, não educados e ensino fundamental), com 27% de analfabetos e 12% de não educados. Esses produtores detêm baixa capacidade de absorção tecnológica, segundo Vieira Filho (2004). Em uma comparação regional, o Nordeste tem o pior desempenho na educação, com 59% dos proprietários analfabetos e não educados, apenas 36% com ensino fundamental e 6% com ensino médio e

⁶O coeficiente de Gini mede o grau de desigualdade na distribuição da renda domiciliar per capita entre os indivíduos. Seu valor pode variar teoricamente desde 0, quando não há desigualdade (as rendas de todos os indivíduos têm o mesmo valor), até 1, quando a desigualdade é máxima (apenas um indivíduo detém toda a renda da sociedade e a renda de todos os outros indivíduos é nula). Série calculada a partir das respostas à Pesquisa Nacional por Amostra de Domicílios (Pnad/IBGE).

superior. Estes números são impressionantes em comparação com outras regiões e mostram uma enorme lacuna entre regiões desenvolvidas e atrasadas.

FIGURA 1 – Percentual de proprietários por níveis de escolaridade e por regiões.



Fonte: Elaboração com base em Vieira Filho (2014).

Com certeza, estas características também são observadas no semiárido. Segundo Aquino e Lacerda (2014), esse ambiente torna atividades tradicionais na região, como a agricultura de sequeiro (que depende somente das precipitações pluviométricas) e a agropecuária extensiva, em empreendimentos de extremo risco. Dada à concentração populacional e sua dependência à essas atividades, períodos prolongados de estiagem acarretam drásticas reduções na produção, aumento da fome, crescimento do desemprego e estímulo a migração para zonas urbanas, o que demanda maior investimento na prestação dos serviços públicos nas grandes cidades. Programas, tais como Bolsa Família, contribuem para amenizar parte desse impacto negativo, mas é preciso planejamento mais específico para acabar com o problema da pobreza de forma estrutural. Políticas assistencialistas são importantes no período de transição, mas investimento em capacidade de absorção de conteúdo tecnológico é fundamental para erradicar com os problemas estruturais.

Diante disso, os formuladores de políticas públicas voltaram sua atenção à essas questões, criando políticas específicas para o enfrentamento dessa condição. O principal instrumento de intervenção governamental, representada pela Superintendência para o Desenvolvimento do Nordeste (SUDENE), é o Fundo Constitucional do Nordeste (FNE), que tem como seu gestor o Banco do Nordeste Brasileiro (BNB). A Constituição Federal de 1988 estabeleceu o FNE e outros fundos constitucionais ⁷ através do artigo 159 e regido por força da Lei N° 7.827, de 27.09.1989, sendo metade dos recursos dos fundos, direcionados à região semiárida.

A delimitação do semiárido brasileiro é regulamentada mediante a Portaria N° 89, assinada em março de 2005 pelo Ministério da Integração Nacional. O semiárido brasileiro constitui toda a área dos estados do Nordeste, incluindo todos os municípios do Norte de Minas Gerais e do Espírito Santo que atendam os seguintes critérios: Precipitação pluviométrica média anual inferior a 800 milímetros; Índice de aridez de até 0,5 calculado pelo balanço hídrico que relaciona as precipitações e a evapotranspiração potencial, no período de 1961 e 1991; e Risco de seca maior que 60%, tomando-se por base o período de 1970 e 1990.

⁷ Os fundos constitucionais foram criados no âmbito da constituição federal de 1988 e são constituídos de 3% da arrecadação do imposto sobre Renda e Proventos de Qualquer Natureza (IR) e do imposto sobre Produtos Industrializados (IPI). Os fundos são: Fundo Constitucional de Financiamento do Nordeste (FNE), Fundo Constitucional de Financiamento do Norte e Fundo Constitucional de Financiamento do Centro-Oeste (FCO).

A importância desse fundo para a região foi mostrada por Da Mata e Resende (2016). No estudo, os autores comparam os municípios incluídos na região semiárida (grupo de tratamento), com municípios semelhantes que não foram incorporados a região na lei de 2005 (grupo de controle). Constatou-se que produtores de baixa renda e com menor porte nas novas localidades semiáridas têm investido o crédito subsidiado para expandir sua capacidade de resiliência às intemperes climáticas regionais. Notou-se ainda que esses empréstimos não se converteram em ganhos em PIB *per capita*, dado o nível de informalidade desse setor na região.

Nessa perspectiva, a região de estudo será a cartografia da área de atuação da SUDENE e não apenas o Nordeste geográfico brasileiro, definido pelo IBGE, no intuito de melhor representar os dois grupos regionais. Embora os municípios do norte de Minas Gerais pertençam a um estado rico, com maior grau de desenvolvimento, não se beneficia desse nível de desenvolvimento, como mostra Iffri (2015). Já o Norte do Espírito Santo foi mantido para garantir a estrutura institucional criada pela atuação da SUDENE na região.

Tabela 2 – Número de municípios abrangidos pela cartografia da SUDENE (2005)

Estados	Número de Municípios		
	Nordeste Semiárido	Nordeste Não-Semiárido	Total
AL	38	64	102
BA	265	152	417
CE	150	34	184
MA	0	217	217
PB	170	53	223
PE	122	63	185
PI	127	96	223
RN	147	20	167
SE	29	46	75
ES	0	28	28
MG	85	83	168
Total	1133	856	1989

Fonte: Portaria MI nº 89 (16.03.2005 - Ministério da Integração Nacional).

Elaboração: Autores.

Esse esforço busca captar o impacto da heterogeneidade espacial nordestina na eficiência das firmas situadas na delimitação proposta, e verificar o quanto as *diferenciações tecnológicas* entre as regiões influencia na eficiência produtiva dos municípios.

O estudo de O'Donnell *et al.* (2008) corrobora com essa percepção, onde firmas situadas em diferentes regiões deparam-se com diferentes oportunidades de produção. As escolhas a serem tomadas são sobre diferentes combinações factíveis de “insumo-produto”, ou seja, a partir de diferentes “conjuntos de tecnologias”. Essas diferenças podem ser explicadas pela disponibilidade e qualidade do estoque de capital físico, humano e financeiro, infraestrutura econômica, dotação de recursos naturais, etc.

2.2. Principais tendências verificadas

Como apresentado na seção anterior, a tendência do setor agrícola do semiárido nordestino é de baixa produtividade. Um cenário dispare quando comparado ao à dinâmica do agronegócio brasileiro. Enquanto em alguns países desenvolvidos discute-se a existência ou não de uma tendência de redução do crescimento da produtividade agrícola, como diversos estudos debatem para o caso do setor agrícola americano (ALSTON e PARDEY, 2009; JAMES

et al., 2009; BALL *et al.*, 2013), países em desenvolvimento, especialmente a China e o Brasil, apresentaram rápidos ganhos de PTF no setor (FUGILE *et al.*, 2008).

Estudos que se detêm a analisar a evolução da PTF e da eficiência agrícola da região semiárida são escassos. A maioria dos trabalhos consideram apenas o Nordeste como um todo. Desconsiderando a elevada heterogeneidade regional, os agricultores fazem as suas escolhas a partir de diferentes “conjuntos de tecnologias”. A literatura econômica sobre o tema mostra que o Nordeste vem a margem desse processo. Vicente (2004), ao analisar a eficiência e a PTF da agricultura brasileira no período de 1970 a 1995, utilizando Análise Envoltória de Dados (DEA), indica que o Nordeste sofreu uma involução tecnológica, o que o fez destoar do aumento da PTF por qual passou o setor agrícola brasileiro. Os fatores principais geradores desta tendência foram a deficiência na infraestrutura local, a dificuldade no acesso e na utilização de insumos modernos e a deterioração das relações de troca na agricultura. A amalgama desses fatores influenciou o desempenho PTF na região.

Araújo e Santos (2009) argumentam que o Nordeste passou por um processo tardio de integração à dinâmica da agricultura brasileira, passando a acompanhar de perto a economia brasileira como um todo, crescendo a taxas médias superiores às nacionais. A base agropecuária da região passou por um processo de diversificação.

Toma-se como exemplo a presença da fruticultura irrigada, especialmente no pólo agroindustrial de Petrolina-Juazeiro no vale sub médio do São Francisco e no moderno pólo de fruticultura do Rio Grande do Norte no Vale do Açu. Da mesma forma, a produção de grãos vem crescendo na direção do MATOPIBA (acrônimo para distinguir os estados do Maranhão, Tocantins, Piauí e Bahia), que se estende aos cerrados nordestino e tocaninense.

Gasques *et al.* (2013) analisam a PTF da economia brasileira no período de 1970-2006. Além disso, apontaram-se que a taxa de crescimento da produtividade dos estados nordestinos, no período 1995/1996 – 2006, apresentou crescimento superior à média nacional, com exceção dos estados da Paraíba e do Rio Grande do Norte.

Fornazier *et al.* (2012), ao estudar o setor agropecuário brasileiro, encontrou fortes evidências do setor ser estruturalmente heterogêneo. Para tanto, compararam-se as diferenças a partir dos dados do Censo Agropecuário. Isso significa que, no caso brasileiro, o setor moderno e produtivo convivia com o atrasado e pobre, bem representados na diferença entre os produtores familiares e não familiares. Entretanto, os autores encontraram que a heterogeneidade não se dá unicamente pelo tipo de produção (familiar ou patronal) ou mesmo pelo tamanho (pequeno ou grande). Há outras características como fatores culturais, geográficos e climáticos. Além do mais, foi constatado que dentro da própria categoria de agricultura familiar havia produção moderna com alta produtividade versus produção atrasada com baixa eficiência técnica produtiva, que produzem praticamente sem excedente monetário.

Vieira Filho (2013) verificou a heterogeneidade estrutural existente na agricultura familiar brasileira em diferentes regiões e constatou que há a coexistência de produtores familiares, modernos e produtivos, com agricultores tradicionais de subsistência com baixo teor tecnológico, sendo estes últimos mais vulneráveis às variabilidades climáticas próprias da região Nordeste. Apresentou-se que a região Nordeste possui a maior heterogeneidade entre os agricultores familiares, já que a maior parte de seus estabelecimentos agropecuários se classificava nesse tipo organização, cerca de 88% do total (IBGE, 2006).

Helfand *et al.* (2015) analisou a PTF e a eficiência técnica com o modelo de fronteira estocástica a partir dos dados dos Censos Agropecuários para os anos 1985, 1995-1996 e 2006. Foi examinado o papel do tamanho da propriedade no acelerado crescimento da PTF agrícola do Brasil nos últimos 20 anos. Para tanto, os estabelecimentos agrícolas foram divididos em classes pelo tamanho e pelas suas respectivas contribuições ao crescimento da PTF nacional. Os resultados específicos para o Nordeste mostraram que a segunda menor classe de estabelecimentos (de 5 a 20 hectares) foi responsável por 18% da produção e 21% dos

estabelecimentos da região. Esta foi a classe que mais evoluiu em eficiência técnica de 7,75%, bem como em relação ao crescimento da PTF (de 4,12%). Não obstante, a maior parte das terras da região, 69% do total, representou maiores fazendas (de 100 a 500 hectares). Esta faixa apresentou as mais reduzidas taxas de crescimentos da PTF.

3. METODOLOGIA

3.1. Meta-fronteira de produção estocástica

A construção da meta-fronteira de produção dá-se em dois estágios. O primeiro estágio é dado pela construção de fronteiras estocásticas regionais. Para tanto, considere que a relação entre insumos e produtos das diversas firmas em diferentes regiões, numa dada indústria, é representada por uma fronteira estocástica regional para cada grupo de firmas. Sendo assim, para cada j -ésimo grupo, existe uma amostra de N_j firmas, produzindo um produto a partir de vários insumos, a fronteira estocástica para estes grupos é definida por:

$$y_{it} = f(X_{it}, \beta^j) \cdot \exp\{v_{it}^j - u_{it}^j\} \quad (1)$$

Onde y_{it} é denotado como a produção observada na firma (neste caso o município) i , no ano t ; X_{it} representa o vetor $1 \times K$ de insumos ou outras variáveis explicativas associadas com a firma; β^j denota o vetor $K \times 1$ parâmetros desconhecidos associados ao grupo j . Considerou-se cada município situado na área que pode ser considerada semiárida como uma firma, divididas em duas regiões, assim j = semiárido e não semiárido. No intuito de testar a hipótese de que as tecnologias usadas nos municípios do semiárido e não-semiárido nordestino são as mesmas, serão estimadas três fronteiras estocásticas, duas fronteiras regionais (Semiárido e Não Semiárido) e uma com todos os municípios.

A composição do termo do erro se dá pelos termos v_{it}^j e u_{it}^j , sendo estes independentes entre si. O primeiro representa as variações por choques aleatórios, erros de medição, etc., que são por hipótese $iid \sim N(0, \sigma^2)$, ou seja, independentes e identicamente distribuídos, com distribuição normal de média zero e variância σ^2 . Já o segundo refere-se à ineficiência técnica própria da firma, que desvia a fronteira de produção para baixo, podendo apresentar diversos tipos de distribuição, como a half-normal, como em Aigner *et al.* (1977), normal truncada (BATTESE *et al.*, 1995), exponencial, gama. Assume-se ainda que: $E(v_{it}) = 0$ (Média zero); $E(v_{it}^2) = \sigma_v^2$ (homocedasticidade); $E(v_{it}v_{jt}) = 0$ para todo $i \neq j$ (não correlacionados); $E(u_{it}^2) = constante$, (homocedasticidade); $E(u_{it}u_{jt}) = 0$ para todo $i \neq j$, (não correlacionados).

As estimações desses parâmetros serão obtidas pelo método da máxima verossimilhança. O resultado da *eficiência técnica* da firma (município) específica é dada pela relação:

$$ET_{it}^j = \frac{y_{it}^j}{f(x_{it}^j, \beta_{it}^j) \cdot \exp\{v_{it}^j\}} = \{ \exp(-u_{it}^j) \}, j = \text{semiárido, não semiárido} \quad (2)$$

Onde $0 \leq ET^j \leq 1$, quando $ET=1$, diz-se que a firma é tecnicamente eficiente.

Para testar a hipótese de que as duas regiões são representadas pela mesma fronteira de produção estocástica, utiliza-se o teste da razão de verossimilhança (LR). Se a hipótese nula for rejeitada, a fronteira de produção agrupada é rejeitada e aceita-se a hipótese alternativa de existência de fronteiras regionais distintas. No entanto, as ET computadas pelas fronteiras regionais estimadas não podem ser comparadas, contudo, adota-se o conceito de função de meta-produção para comparação das diferenças de performances entre as duas regiões de estudo.

Parte-se da hipótese de que os municípios estão divididos em dois grupos regionais (Semiárido e Não-Semiárido) e possuem diferentes *conjuntos de produção*, dados seus

diferentes níveis de acesso aos estoques de capital físico, humano e financeiro, como nas diferentes dotações de recursos naturais, disparidade na qualidade do solo, clima, acesso à recursos energéticos, infraestrutura, etc. Assim, a MF possibilita a comparação das TE das diferentes fronteiras estimadas. A função de MF de produção toma a seguinte forma:

$$y_{it}^* = f(X_{it}, \beta^*) \equiv \exp\{X'_{it}\beta^*\} \quad (3)$$

Em que y_{it}^* indica o produto da meta-fronteira e β^* indica o vetor de parâmetros para a MF de produção que satisfaça a seguinte restrição:

$$X'_{it}\beta^* \geq X'_{it}\beta^j, \text{ para } j = \text{semiárido, não semiárido} \quad (4)$$

Essa restrição garante que nenhuma fronteira regional seja maior que a meta-fronteira. O'Donnell et al. (2008) apresenta duas formas para o cálculo β^* . O primeiro método é estimar uma meta-fronteira estocástica, usando os produtos, insumos de todas as firmas, independentemente da região e em todos os períodos, calculando uma meta-fronteira estocástica. Este método não garante que a meta-fronteira estimada seja a envoltória das fronteiras regionais, dada a possibilidade de haver problema de má especificação. O segundo método, proposto por Battese et al (2004) e O'Donnell et al (2008) considera que, dado os parâmetros estimados $\hat{\beta}^j$ das fronteiras regionais no primeiro estágio, o segundo estágio para a estimação da MF se dá na estimação de β^* pela solução do seguinte problema de programação linear:

$$\min_{\beta} L \equiv \sum_{t=1}^T \sum_{i=1}^N |\ln f(X_{it}, \beta^*) - \ln f(X_{it}, \hat{\beta}^j)| \quad (5)$$

$$\text{s.t. } \ln f(X_{it}, \beta^*) \geq \ln f(X_{it}, \hat{\beta}^j), \text{ para todo } i \text{ e } t \quad (6)$$

Onde $\hat{\beta}^j$ é o vetor dos coeficientes estimados das j-ésimas fronteiras estocásticas. No caso destes parâmetros estimados para as fronteiras estocásticas nas diferentes regiões ($\hat{\beta}^j$, $j = \text{semiárido e não-semiárido}$) serem fixos para todo o problema, e sendo a função $f(\cdot)$ log linear nos parâmetros, como foi assumido em (1), o problema linear toma a seguinte forma:

$$\min_{\beta} \bar{X}\beta^* \quad (7)$$

$$\text{s.t. } X'_{it}\beta^* \geq X'_{it}\beta^j, \text{ para todo } j, i \text{ e } t \quad (8)$$

Onde $\bar{X}'$ é o vetor coluna das médias dos elementos dos x_{it} -vetores para todas as observações no conjunto de dados. Com o problema de programação linear resolvido, tem-se o vetor β^* definido, rescreve-se o produto observado da firma (município) i , na região j no ano t , pela seguinte decomposição:

$$y_{it}^j = \exp(-u_{it}^j) \times \frac{f(X_{it}^j, \beta^j)}{f(X_{it}^j, \beta^*)} \times f(X_{it}^j, \beta^*) \exp(v_{it}^j) \quad (9)$$

$$MTR_{it}^j = \frac{f(X_{it}^j, \beta^j)}{f(X_{it}^j, \beta^*)}, \text{ onde } 0 \leq MTR_{it}^j \leq 1 \quad (10)$$

O primeiro termo do lado direito da equação (9) representa a eficiência técnica (ET) relativa a fronteira estocástica da j-ésima região. O segundo termo é nomeado de razão de meta-tecnologia (MTR)), que representa a diferença entre a tecnologia disponível para a região j , relativa a melhor tecnologia disponível para a indústria como um todo, e exposta na equação (10). A eficiência técnica relativa a meta-fronteira estocástica é definida de uma forma análoga a equação (11):

$$ET_{it}^{*j} = \frac{y_{it}^j}{f(X_{it}^j, \beta^*) \exp(v_{it}^j)} = ET_{it}^j \times MTR_{it}^j \quad (11)$$

3.2. Base de dados

Os dados utilizados correspondem às 1989 informações dos municípios que compõem a área de atuação da SUDENE. As fontes de informações baseiam-se nos Censos Agropecuários do IBGE de 1970 a 2006, além do sistema de contas regionais, também disponibilizadas pela

plataforma do IPEADATA, em uma série temporal de 1970, 1975, 1980, 1985, 1995/1996 e 2006. Os municípios não existentes nos períodos do censo, ou que possuíam informações omitidas para alguma variável em algum dos anos de análises, foram excluídos da amostra, de forma a manter os dados em painel balanceado. Dos 1989 municípios iniciais, a amostra final contabilizou 1268 municípios que continuaram nesta análise, sendo 740 pertencentes ao semiárido e 528 ao não-semiárido.

A variável escolhida para representar o produto foi o PIB agrícola municipal, disponibilizado pelo IPEADATA e deflacionado a preços de 2000 pelo deflator implícito do PIB. Para a mensuração dos fatores de produção terra, trabalho e despesas com insumos, a coleta de dados dependeu exclusivamente dos dados disponibilizados pelos censos agropecuários. O fator terra foi considerado como a área total das terras plantadas com lavouras permanentes e temporárias. Não se contabilizou aí áreas plantadas com pastagens, por não serem considerados como fator escasso para agropecuária como apresenta Bragagnoli (2012).

Para medir a força de trabalho, optou-se por utilizar os dados referentes ao pessoal ocupado na agropecuária, que incluí tanto proprietários, quanto trabalhadores contratados. Já as despesas com insumos foram retiradas das despesas declaradas em fertilizantes, defensivos, sementes, mudas, energia, dentre outros. Estes valores estão deflacionados pelo deflator implícito do PIB nacional. Para a variável estoque de capital, utilizou-se o valor dos bens dos agricultores, disponibilizados pelo censo agropecuário, incluindo as construções rurais, máquinas e equipamentos. Contudo, essa variável não foi disponibilizada no censo agropecuário de 1995/1996, o que obrigou a utilização de uma proxy para este período em especial, tal como proposto por Bragagnoli (2012). O método consistiu de partir dos dados disponíveis em 1985, dos quais se acumularam os valores dos investimentos anuais até 1995, aplicando ainda uma depreciação de 6%, de acordo com os valores apresentados em Gasques *et al.* (2013). Esta variável foi deflacionada em Reais (R\$) de 2000 pelo deflator implícito da formação bruta de capital, disponibilizado pelo IPEADATA (2016).

3.3. Especificação do modelo empírico

Para o cálculo da meta-fronteira de produção, deve-se primeiramente estimar de fronteiras regionais por fronteira estocástica. Utiliza-se o método proposto por Aigner, Lovell e Schmidt (1977) e Meusen e Broeck (1977) de fronteira estocástica de produção, o qual foi aprimorado por Pitt e Lee (1981) e Schmidt e Sickles (1984), permitindo, dessa maneira, a modelagem de dados em painel, que incorporou o componente de ineficiência técnica de produção, junto às contribuições relativas a Battese e Coelli (1992;1995).

Nesse sentido, inicialmente definiu-se a forma funcional da fronteira estocástica de produção. Existem diferentes formas funcionais que dão origem a diferentes modelagens econométricas na análise de fronteiras estocásticas. As formas funcionais mais comuns na literatura são a Cobb-Douglas (CD) e a Translog (TL). Good *et al.* (1993) e Ahmad e Barvo-Ureta (1996) argumentam que a mensuração da ET é robusta para ambas formas funcionais.

Dessa forma, utilizar-se-á a forma funcional CD pela sua simplicidade e dado o objetivo de estudo. A escolha realizada deu-se pelos desafios inerentes à estimação de fronteiras regionais com os mesmos fatores de produção, pois as diversas formas funcionais testadas tinham indícios de má-especificação para pelo menos uma das fronteiras estimadas. Sendo assim, optou-se por uma abordagem mais simples, ainda assim robusta, como salientado. A fronteira produção estocástica é especificada da seguinte forma:

$$\ln Y_{it} = \beta_1 + \beta_2 \ln L_{it} + \beta_3 \ln T_{it} + \beta_4 \ln K_{it} + \beta_5 \ln I_{it} + \beta_6 t + v_{it} - u_{it} \quad (12)$$

Onde: Y_{it} = Produto agropecuário do município i no período t ; L_{it} = Fator de produção terra do município i no período t ; T_{it} = trabalho por município i no período t ; K_{it} = estoque de capital

por município i no período t ; I_{it} = despesas com Insumos por município i no período t ; t = tendência linear; v_{it} = distúrbios aleatórios da função de produção que por hipótese $iid \sim N(0, \sigma^2)$; u_{it} = ineficiência técnica da produção que por hipótese assume-se que $iid \sim HN(0, \sigma^2)$.

A estimação dos parâmetros da equação (12) é feita pelo método da máxima verossimilhança, permitindo o cálculo das magnitudes das eficiências técnicas para cada um dos municípios da amostra e, conseqüentemente, possibilitando a estimação das fronteiras regionais (semiárido e não semiárido), quanto a fronteira agrupada. Após as estimações dos modelos, foram realizados os respectivos testes de forma funcional para escolha do melhor modelo, o de ineficiência técnica, o teste de ausência de progresso técnico, como também o de existência de fronteiras regionais.

4. Análise e resultados

Procura-se dividir a análise em três partes. A primeira seção contempla a análise das estatísticas descritivas. A segunda encontra-se o primeiro estágio da estimação da meta-fronteira, que é constituído pela estimação das fronteiras regionais na abordagem da fronteira estocástica, bem como seus resultados. Por fim, tem-se o segundo estágio, no qual se estima a meta-fronteira de produção por técnicas de programação linear, analisando os seus resultados.

4.1. Estatísticas descritivas

As estatísticas descritivas estão expostas na Tabela 3. Conforme os dados, no que tange às médias, aos desvios padrões e às amplitudes, observam-se variações significativas entre as regiões estudadas, o semiárido e os demais municípios. No âmbito inter-regional, as duas regiões possuem grandes disparidades, principalmente, entre os valores máximos e mínimos, o que permite inferir que esse desequilíbrio entre os usos dos fatores de produção impacta na eficiência dos municípios na produção agrícola.

O produto agrícola do semiárido é menor, em média, do que no não-semiárido. Todavia, apresenta um menor desvio padrão e menor amplitude, representando que os municípios do semiárido são mais homogêneos que os do não-semiárido. No que tange aos fatores de produção, o semiárido consome, em média, menos em quase todos insumos, com exceção do fator de produção *terra*. Isso pode indicar uma relação dessa região diferenciada com a terra.

O não-semiárido gasta, em média, quase o dobro de insumos quando comparado ao semiárido, e possui o fator *capital médio* significativamente mais elevado. Isso significa que o não-semiárido possui um quantitativo superior de maquinaria, veículos e construções. No que se concerne às despesas com insumos, o não-semiárido consome, em média, uma maior quantidade de energia, de mudas, de serviços veterinários, dentre outros, quando comparado ao semiárido.

Tabela 3 – Estatísticas descritiva

Variáveis	Região	Média	Desvio Padrão	Min	Max
Produto (mil reais)	Semiárido	8273,20	11208,94	38,51	216760,36
	Não-Semiárido	12375,00	18490,09	31,00	360224,00
Área (hectare)	Semiárido	8768,55	10843,26	60,22	265468,37
	Não-Semiárido	8284,13	10124,78	1,76	176136,44
Trabalho (Pessoas ocupadas)	Semiárido	6210,00	5302,72	37,00	39477,00
	Não-Semiárido	6256,00	6077,623	18,00	67187,00
Capital (Mil reais)	Semiárido	102031,00	280192,44	234,00	9905797,00
	Não-Semiárido	150493,00	391280,5	159,00	7812573,00

Despesas com	Semiárido	3933,40	13480,83	53,20	691557,50
Insumos (Mil reais)	Não-Semiárido	6989,10	23673,91	0,10	865807,70

Fonte: elaborada pelos autores com base os resultados da pesquisa.

4.2. Estimação das fronteiras de produção e testes de especificação

Os resultados das estimações⁹ estão presentes na Tabela 4. Vale salientar que, para todas as fronteiras, foi realizado o teste de hipótese pelo teste razão de verossimilhança generalizada, no intuito de verificar se o termo do erro correspondente a ineficiência seria inexistente. Rejeitou-se a hipótese nula de inexistência da eficiência técnica em todas as fronteiras a um nível de significância de menos de 1%.

Nas três fronteiras estimadas, todos os parâmetros são altamente significativos a um nível de 1%, implicando em uma função de produção bem-comportada, com quase todos os parâmetros com sinais esperados. Chamou-se a atenção o quão semelhante as duas fronteiras regionais foram, sendo as mesmas variáveis, impactando positivamente ou negativamente na produção agrícola em ambas as regiões.

Os fatores de produção *terra*, *trabalho* e *estoque de capital* impactam positivamente no produto agrícola em ambas fronteiras regionais como esperado. Existiam diferenças regionais nos impactos dessas variáveis na produção, sendo que o semiárido mantém-se a retaguarda comparativamente as outras regiões em relação ao fator terra. Porém, a influência do trabalho e do capital é superior no semiárido, o que sinaliza que investimentos em educação e em tecnologia são essenciais para transpor os desafios climáticos e produtivos da região.

O fator *trabalho* impacta positivamente na produção em ambas fronteiras regionais, apontando o quão importante é o investimento em educação para o Nordeste como um todo, no sentido de potencializar esses ganhos. Esta assertiva é confirmada por Bacelar e Santos (2009), que percebem que o Nordeste necessita de melhoras na educação para alcançar o desempenho nacional, como apresentou-se na figura 1.

É perceptível que o fator *capital* influencia positivamente na produção agropecuária e é ligeiramente maior no semiárido, corroborando com os resultados de Gasques (2013). Dessa forma, a modernização do maquinário, que substituisse tratores de menor potência pelo uso de veículos de médio e grande porte, acarretaria ganhos na produção do setor. Daí a importância da política de juros subsidiados proporcionados pelo FNE, com enfoque no semiárido, na ampliação do estoque de capital.

Tabela 4 – Fronteiras Estocásticas Estimadas

Variáveis	Parâmetros	Nordeste	Não-Semiárido	Semiárido
(intercepto)	β_1	3,0*** (0,12)	3,24*** (0,16)	2,55*** (0,15)
L_{it}	β_2	0,21*** (0,01)	0,24*** (0,02)	0,20*** (0,02)
T_{it}	β_3	0,25*** (0,01)	0,23*** (0,02)	0,27*** (0,02)
K_{it}	β_4	0,23*** (0,01)	0,22*** (0,01)	0,24*** (0,01)

⁹ Para estimação dos parâmetros das fronteiras de produção, utilizou-se o software estatístico *R-Project*, especificamente o pacote *frontier* (COELLI; HENNINGSEN, 2013).

I_{it}	β_5	-0,15*** (0,01)	-0,12*** (0,02)	-0,15*** (0,01)
	β_6	-0,03*** (0,01)	-0,05*** (0,00)	-0,02*** (0,00)
t	σ^2	1,13*** (0,05)	0,90*** (0,06)	1,03*** (0,06)
	γ	0,73*** (0,01)	0,66*** (0,02)	0,71*** (0,02)
	σ_u^2	0,82*** (0,05)	0,60*** (0,06)	0,72*** (0,06)
	σ_v^2	0,30*** (0,00)	0,30*** (0,00)	0,30*** (0,00)
	σ	1,06*** (0,02)	0,95*** (0,03)	1,01*** (0,03)
	σ_u	0,91*** (0,03)	0,77*** (0,03)	0,85*** (0,03)
	σ_v	0,55*** (0,00)	0,55*** (0,00)	0,55*** (0,03)
	λ^2	2,74*** (0,17)	1,98*** (0,21)	2,41*** (0,22)
	λ	1,67*** (0,05)	1,41*** (0,07)	1,55*** (0,07)
	ET Média	0,52	0,58	0,54
	log-verossimilhança	-7359,52	-3011,32	-4259,06

Fonte: Elaborada pelos autores com base nos resultados da pesquisa.

*A codificação para a significância dos coeficientes estimados:

‘***’ significante a um nível de 1%.

O fator de produção *terra* influencia positivamente a produção em ambas fronteiras regionais, sendo mais significativa no não-semiárido nordestino. O trabalho de Helfand *et al.* (2015) discute a estrutura fundiária do Nordeste. Segundo os autores, os estabelecimentos mais eficientes e produtivos da região Nordeste são as pequenas propriedades; todavia, 69% das terras estão em propriedade dos estabelecimentos de grande porte, os menos produtivos da Região. A incorporação de novas áreas para os produtores mais eficientes e produtivos da região é bastante restrita, podendo caracterizar-se um entrave ao aumento da produtividade via aumento da área cultivada desses pequenos produtores eficientes.

Já as variáveis *despesas com insumos* e a *tendência temporal* possuem uma relação inversa com a produção em ambas fronteiras regionais. No que tange *despesas com insumos*, há uma indicação de causalidades perante às condições próprias da região, que tornam estes gastos forçosos ao processo produtivo, pois as três fronteiras apresentaram a mesma relação, sendo ainda mais forçosa para o nordeste semiárido. Nesse ponto, quanto maiores são os gastos com insumos, menor seria o retorno produtivo, ou seja, o uso intensivo de insumo não representa aumento de produção. Esse é um resultado curioso, mas pode estar associado à baixa capacidade de absorção tecnológica dos agentes, já que estes são agricultores com baixo nível de escolaridade.

Isso se dá pela região demandar intensamente adubos, defensivos, gastos com energia e água, para manter uma área irrigada, por exemplo, como também pode-se inferir a possibilidade da má utilização desses mesmos fatores, dada a baixa qualificação da mão-de-obra. Vicente (2004) aponta como possível causa para o baixo dinamismo da região o caso da infraestrutura ineficiente e o limitado acesso ao crédito assimétrico na região, inviabilizando a aquisição de insumos modernos. Há indícios de uma reversão dessa involução técnica do Nordeste a partir

de 1985, dada uma maior integração do Nordeste com o mercado interno brasileiro e o advento do FNE. Porém, não se percebe esta tendência nos resultados aqui apresentados.

A variável *tendência linear* foi significativa para ambas fronteiras e indica uma diminuição da dinamicidade do setor agrícola da região no decorrer da série histórica, corroborando com os resultados encontrados por Vicente (2004). Essa tendência pode ser um indicativo de que o Nordeste brasileiro ficou à margem do processo de modernização agrícola, que tornou a agricultura do país competitiva no cenário internacional, como mostra O'Donnell *et al.* (2008) e Gasques (2014), tornando o Nordeste uma região de baixo dinamismo no setor agrícola, como demonstra Vieira Filho (2013). Percebe-se ainda que essa tendência é ainda mais intensa no não-semiárido nordestino, o que é um resultado inesperado.

A média da *eficiência técnica* varia entre as fronteiras regionais. O não-semiárido apresenta ET média de 58,07% e o semiárido possui o escore médio de eficiência de 53,8%. Esta diferença não implica que os municípios do semiárido possuem desempenhos piores que os do não-semiárido, pois a comparação das duas fronteiras só fará sentido se elas não compartilharem da mesma tecnologia de produção. No entanto, se as duas regiões comungam da mesma tecnologia, deve-se usar o resultado dos municípios agrupados do Nordeste.

4.3 Análise dos Testes de Hipóteses

Após as estimações dos modelos, foram realizados os respectivos testes de forma funcional para escolha do melhor modelo, o de ineficiência técnica, o teste de ausência de progresso técnico, como também o de existência de fronteiras regionais. Como discutido, por problemas de convergência, optou-se pela forma funcional Cobb-Douglas, pois mostrou-se a mais adequada para às três fronteiras estimadas, não havendo indícios de má-especificação.

Escolhida a forma funcional, realizou-se o teste de ausência de ineficiência técnica ao modelo, caso $\gamma = 0$, não se deve utilizar o Modelo de Fronteira Estocástica, ou seja, o modelo não capta o efeito ineficiência. Nesse caso, é indicado a utilização do método de mínimos quadrados ordinários. A partir dos respectivos valores de log máxima verossimilhança das estimações de cada fronteira, os valores λ estão dispostos na Tabela 5. Tem-se os resultados do teste para as três fronteiras, que superam seus respectivos valores críticos e significância de 5% na tabela de Kodde e Palm (1986), dados seus graus de liberdades. Portanto, rejeita-se H_0 e aceita-se a hipótese H_1 , confirmando a presença de progresso técnico para às três fronteiras apresentadas.

Posteriormente, testou-se a ausência de progresso técnico. De acordo com o teste descrito anteriormente, estimou-se o modelo na forma funcional Cobb-Douglas e na forma com ausência de progresso técnico. De acordo com os resultados, rejeitou-se a hipótese nula e aceitou-se a hipótese alternativa, validando o uso do modelo de fronteira estocástica.

Tabela 5 – Prova de Razão Verossimilhança dos parâmetros das fronteiras de Produção

Prova	Região	Hipótese nula	Graus de Liberdade	Valor de λ	Valor Crítico	Decisão (valor de 5%)
Inexistência de Eficiência Técnica	Fronteira Agrupada	$H_0: \gamma = 0$	1,00	1937,60	6,64	Rejeita H_0
	Fronteira Semiárido	$H_0: \gamma = 0$	1,00	873,56	6,64	Rejeita H_0
	Fronteira Não-Semiárido	$H_0: \gamma = 0$	1,00	617,70	6,64	Rejeita H_0
	Fronteira Agrupada	$H_0: \beta_6 = 0$	1,00	37,80	6,64	Rejeita H_0

Ausência de Progresso Técnico	Fronteira Semiárido	$H_0: \beta_6 = 0$	1,00	7,82	6,64	Rejeita H_0
	Fronteira Não-Semiárido	$H_0: \beta_6 = 0$	1,00	41,12	6,64	Rejeita H_0
Existência de duas Fronteiras Regionais	Não se aplica	H_0 : Fronteira Agrupada	8,00	178,00	25,37	Rejeita H_0

Fonte: elaboração própria.

λ : prova estatística da razão de máximo verossimilhança em que $\lambda = -2[LL(H_0) - LL(H_1)] \sim \chi^2$, com graus de liberdade iguais ao número de restrições independentes.

Por fim, realizou-se o teste para validar a existência da diferenciação entre as possibilidades produtivas entre as duas regiões ou não. Conforme os resultados, foi rejeitada a hipótese nula e aceita a hipótese alternativa, validando que os municípios pertencentes ao semiárido ou ao não semiárido empregam diferentes tecnologias de produção. Logo, pode-se utilizar o quadro teórico da meta-fronteira de produção, a fim de comparar a performance de eficiências entre as diferentes regiões.

4.3. Análise das eficiências técnicas e razão de meta-fronteira

Para a estimação dos parâmetros da meta-fronteira tecnológica, foi resolvido o problema de programação linear apresentado, utilizando o software estatístico *Shazam*. Os parâmetros encontrados para a meta-fronteira correspondem aos apresentados na fronteira regional dos municípios situados no não-semiárido, estimada pelo modelo de fronteira estocástica. Este resultado indica que a tecnologia de produção empregada na agropecuária nestes municípios é mais avançada do que na região semiárida. O Resultado aponta que a região não-semiárida define tecnologia potencial disponível para todos os municípios do Nordeste.

A razão de meta-tecnologia (MTR), definida na equação (10), é para todos os municípios do não-semiárido igual a 1, pois eles situam-se na meta-fronteira de produção. A MTR média dos municípios do semiárido é por volta de 64% com um desvio padrão de 4%. Esse resultado indica que, em média, os municípios do semiárido produzem no máximo 68% do máximo que pode ser produzido na região não-semiárida ao utilizar os mesmos fatores de produção e tecnologias disponíveis no não-semiárido.

Já a ET referente a meta-fronteira de produção é chamada de ET^* e foi estimada para todos os municípios do semiárido de acordo com a equação (11). Os municípios do não-semiárido, por serem pertencentes a meta-fronteira de produção tem a sua ET^* equivalente à calculada sobre a fronteira regional. A Tabela 6 resume os resultados referente às eficiências e a MRT.

Tabela 6 - Estatística Descritiva e Distribuição de frequência das EF e MTR estimadas

Estatísticas	$ET_{Semiárido}$	$MTR_{Semiárido}$	$ET^*_{semiárido}$	$\frac{ET_{não-semiárido}}{ET^*_{não-semiárido}}$
Média	0,54	0,64	0,35	0,58
Desvio Padrão	0,20	0,04	0,13	0,19
Mín.	0,04	0,54	0,03	0,07
Máx.	0,96	0,76	0,67	0,97
Níveis				
<0.6	63,7%	19,2%	95,3%	51,5%
0.6-0.7	13,4%	74,9%	4,7%	17,4%

0.7-0.8	9,9%	5,9%	0,0%	14,4%
0.8-0.9	8,5%	0,0%	0,0%	13,7%
0.9-1	4,5%	0,0%	0,0%	3,1%

Fonte: Elaborada pelos autores com base nos resultados da pesquisa.

O semiárido apresenta uma ET* 35%, tendo como referência à MF entre 1970 e 2006. O município mais ineficiente da região possui um escore de ET* de 3%, enquanto o município mais eficiente possui um escore de 67%. Já a região não-semiárida tem um escore de ET* média de 58%, substancialmente superior à outra região de análise, onde a sua amplitude é tão grande quanto a apresentada no semiárido, possuindo valor mínimo na grande de 7% e máximo de 97%.

Quando se compara apenas as eficiências regionais, tem-se a impressão enganosa de que a diferença entre as duas regiões no que tange a ET não seria tão gritante. Ao comparar as distribuições de frequências das $ET_{\text{não-semiárido}}^*$ e $ET_{\text{semiárido}}^*$, ainda na Tabela 6, percebe-se que todas as observações no semiárido encontra-se entre 0.0 e 0.7, enquanto o não-semiárido se distribui de maneira mais homogênea, quando comparada. Essas conclusões não implicam que a situação do não semiárido seja das melhores. Nesta região, a maioria dos municípios apresentam escores inferiores a 60%, tornando ainda mais grave a situação no semiárido nordestino, tendo em vista a dependência das camadas mais pobres da população com a atividade agropecuária.

Quando se agrupam os municípios em estados, percebe-se uma uniformização da MRT no período de análise, como disposto na Tabela 7. Ao comparar as ETs* percebe-se que semiárido piauiense apresenta pior performance (27,2%), seguindo pelo semiárido paraibano (30,5%), potiguar (31,3%) e sergipano (31,9%). Os maiores escores são apresentados pelo semiárido cearense (37,8%), acompanhado de perto do semiárido pernambucano (37,6%) e mineiro (37,6%), mesmo assim com percentuais bastante inferiores.

Tabela 7 – Amostras sub-regionais das ET e MRT médias

Região Semiárido (1970-2006)	ET	MRT	ET*	Contagem
AL	0,49	0,64	0,32	204
BA	0,58	0,64	0,37	1224
CE	0,59	0,64	0,38	702
MG	0,58	0,64	0,37	48
PB	0,48	0,64	0,31	504
PE	0,58	0,64	0,38	624
PI	0,43	0,64	0,27	288
RN	0,49	0,64	0,31	672
SE	0,50	0,64	0,32	174
Região Não-semiárida (1970-2006)	ET=ET*		Contagem	
AL	0,59		336	
BA	0,61		702	
CE	0,52		126	
ES	0,86		90	
MA	0,62		744	
MG	0,60		30	
PB	0,49		186	
PE	0,69		306	

PI	0,39	300
RN	0,48	96
SE	0,48	252

Fonte: Elaborada pelos autores com base nos resultados da pesquisa.

O não-semiárido dos estados do Nordeste apresenta melhor performance na eficiência técnica com relação a meta-fronteira, se comparados ao semiárido. Percebe-se, nesse conjunto de Estados, uma disparidade nos municípios pertencentes ao Norte do Espírito Santo (86%), dado a distância da sua performance técnica com as médias dos demais estados. Esse resultado pode apontar uma dissociação do Norte do Espírito Santo no que tange o perfil dos municípios que compõem a área de atuação da SUDENE.

Percebe-se que a opção de manter o norte de Minas Gerais na amostra de municípios pertencentes à cartografia do Nordeste considerada pela SUDENE foi acertada. Identifica-se a homogeneidade da região, podendo-se inferir que, mesmo pertencendo ao Estado de Minas Gerais, um estado mais desenvolvido, que possui maior arrecadação que os Estados do Nordeste e, conseqüentemente, maior capacidade de ofertar bens e serviços públicos, ainda assim, não é capaz de diferenciar seus municípios das condições no que tange a eficiência produtiva. Esse resultado aproxima-se de Iffri (2015).

Ainda nas melhores práticas, é possível citar o não-semiárido pernambucano (69%). O semiárido delimitado pelo estado com o menor escore é o Piauí, o que mostra o quão problemático é a agropecuária do estado, seguido pelo não semiárido paraibano (30,5%), potiguar (31.3%) e sergipano (32%), novamente, tornando estes estados com as piores práticas nos dois estratos. Vale ressaltar que esses dados não captam a pungente produção de grãos que vem se instalando no Piauí, dado o horizonte temporal utilizado, o que pode não captar as transformações recentes que estão ocorrendo no estado, bem como seus impactos na eficiência técnicas de seus municípios.

CONSIDERAÇÕES FINAIS

O objetivo deste estudo foi o de analisar as diferenças tecnológicas da agropecuária entre os municípios situados no semiárido e no não-semiárido do Nordeste brasileiro. Inicialmente, definiu-se a área de atuação da SUDENE e não apenas o Nordeste geográfico brasileiro, definido pelo IBGE, no intuito de melhor representar os dois grupos regionais. O modelo utilizado foi de meta-fronteira proposto por Battese et al (2004) e O'Donnell et al (2008) que possibilitou as estimações das eficiências técnicas (ET e ET*) e da razão de meta-fronteira (MRT) dos municípios inseridos no semiárido e não-semiárido nordestino.

A partir desses procedimentos metodológicos, pôde-se afirmar que o semiárido e o não semiárido nordestino não compartilham da mesma tecnologia de produção, aceitando a assertiva assumida como hipótese. Para tanto, foram utilizados dados do Censo Agropecuário e das Contas Nacionais de 1970 a 2006. Pelas fronteiras de produção estimadas, indica-se que as políticas públicas da agropecuária da região devem levar em conta a importância do investimento em qualificação da mão-de-obra, a necessidade de direcionamento do crédito subsidiado disponibilizados pelo FNE para o aumento e modernização do estoque de capital e aquisição de insumos modernos. Outro ponto importante é fortalecer o papel da Embrapa na região como difusora de conhecimento e de novas tecnologias, melhorando os níveis de escolaridade da população e os serviços de extensão rural com capilaridade local.

Buscou-se atentar à análise da questão da produção agrícola do Nordeste, a saber, que diferenças regionais, próprias do semiárido Nordestino, impactam nas práticas agrícolas dos municípios, impondo-lhes diferentes combinações factíveis de *insumo-produto*, a partir de diferentes *conjuntos de tecnologias*. Essa constatação não implica que a organização produtiva

da agropecuária no não-semiárido seja moderna, aproximando-a das regiões mais dinâmica do país.

Na verdade, os problemas no setor agrícola do Nordeste, identificados extensamente pela literatura econômica, impactam de maneira ainda mais severa no semiárido nordestino. Numa análise em âmbito estadual, os municípios situados no Espírito Santo se sobressaíram quando comparados aos demais agrupamentos estaduais dos municípios situados no não-semiárido. No entanto, o Norte de Minas Gerais não se destaca das demais, mesmo pertencendo a um Estado com um maior grau de desenvolvimento. Os estados que apresentaram as piores práticas para ambos as regiões do Nordeste foram Piauí, Paraíba, Rio Grande do Norte e Sergipe.

Os resultados indicaram que, quando comparadas as duas regiões, o não-semiárido possui a maior eficiência técnica média com referência a meta-fronteira, de 58% contra 35% do semiárido. Este resultado sugere que, mesmo lançando mão de todas as possibilidades produtivas da região Nordeste, o não-semiárido tem uma prática bem aquém do ideal, pois estão utilizando fatores de produção bem abaixo de suas reais potencialidades. Por fim, estudo não discute as razões dessas disparidades regionais no que concerne a MTR e as eficiências técnicas entre ambas fronteiras regionais e a meta-fronteira. Este tema é digno de aprofundamentos em investigações futuras.

Referências Bibliográficas

- AB'SABER, A. Sertões e Sertanejos: uma Geografia Humana Sofrida. Estudos Avancados, v. 13, p. 7–59, 1999.
- AHMAD, M.; BRAVO-URETA, B.E. Technical efficiency measures for dairy farms using panel data: a comparison of alternative model specifications. Journal of Productivity Analysis, 7, p. 399–415, 1997.
- AIGNER, D. J.; LOVELL, C. A. K. SCHMIDT, P. Formulation and estimation of stochastic frontier production functions models. Journal of Econometrics, v.6, p. 21-37, 1977.
- ALSTON, Julian M. et al. Persistence pays: US agricultural productivity growth and the benefits from public R&D spending. Springer Science & Business Media, 2009.
- ALVES, E.; ROCHA, D. P. Ganhar tempo é possível? In: GASQUES, J. G.; VIEIRA FILHO, J. E. R.; NAVARRO, Z. (Orgs.). A agricultura brasileira: desempenho, desafios e perspectivas. Brasília: IPEA. 2010. Cap. 11, p. 275-290.
- AQUINO, J. R. de; LACERDA, M. A. D. de. Magnitude e condições de reprodução econômica dos agricultores familiares pobres no semiárido brasileiro: evidências a partir do Rio Grande do Norte. Revista de Economia e Sociologia Rural, v. 52, p. 167-188, 2014.
- ARAÚJO, T. B.; SANTOS, V. Desigualdades regionais e nordeste em formação econômica do Brasil. In Araújo, Tarcísio; Vianna, Salvador & Macambira, Júnior. (Org). 50 anos de Formação Econômica do Brasil - Ensaio sobre a obra clássica de Celso Furtado. Rio de Janeiro, IPEA, 2009.
- BALL, E.; SCHIMMELPFENNIG, D.; WANG, S. L. Is US Agricultural Productivity Growth Slowing?. Applied Economic Perspectives and Policy, v. 35, n. 3, p. 435-450, 2013.
- BARBIERI, A.F., DOMINGUES, E., QUEIROZ, B. L., RUIZ, R. M., RIGOTTI, J. I., CARVALHO, J. A., RESENDE, M. F. Climate change and population migration in Brazil's Northeast: scenarios for 2025–2050. Population and environment, v. 31, n. 5, p. 344-370, 2010.
- BATTESE, G.E.; COELLI, T.J. A model for technical inefficiency effects in stochastic frontier production functions for panel data. Empirical Economics, 20, p.325-332, 1995.
- BATTESE, G.E.; COELLI, T.J. Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. Journal of Productivity Analysis 3:153–169. 1992.

BATTESE, G.E.; RAO, D.S.P., O'DONNELL, C. J. A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. *Journal of Productivity Analysis*, 21, p. 91–103. 2004.

BRAGAGNOLO, C. Produtividade, crescimento e ciclos econômicos na agricultura brasileira. Piracicaba – SP, 2012. 168 p. Tese (Doutorado) – Escola Superior de Agricultura “Luiz de Queiroz” – ESALQ/USP.

COELLI, T. J.; HENNINGSEN, A. frontier: Stochastic Frontier Analysis. R package version 1.1-0. <http://CRAN.R-Project.org/package=frontier>. 2013

DA MATA, D.; RESENDE, G. Changing the Climate for Banking: The Economic Effects of Credit in a Climate-Vulnerable Area. 2015. Disponível em: <http://lacer.lacea.org/handle/123456789/52983>

DAVIS, M. Holocaustos Coloniais. Record, Rio de Janeiro. 2002

FORNAZIER, A.; VIEIRA FILHO, J. E. R. Heterogeneidade estrutural no setor agropecuário brasileiro: evidências a partir do censo agropecuário de 2006. Texto para Discussão (IPEA. Brasília), v. 1, p. 7-30, 2012.

FUGLIE, K.O.; WANG, S.L. New Evidence Points to Robust but Uneven Productivity Growth in Global Agriculture. Amber Waves Feature, Economic Research Service (ERS), U.S. Department of Agriculture. September 20, 2012. Disponível em: <http://ers.usda.gov/amber-waves/2012-september/global-agriculture.aspx>.

GASQUES, J.; BACCHI, M.; FUGLIE, K. O.; WANG, S. L.; BALL, V. E. Total factor productivity in Brazilian agriculture. *Productivity Growth in Agriculture: An International Perspective*. CAB International, Oxfordshire, UK, p. 145-162, 2012. *Perspective* (Eds. Fuglie, et al.), Capítulo 7, p. 145-162.

GASQUES, J.G. Total Factor Productivity in Agriculture: Brazil and Selected States. Productivity Note, Ministry of Agriculture, Livestock, and Food Supply (Ministério da Agricultura, Pecuária e Abastecimento - MAPA). Fevereiro, 2014.

GOOD, D.H.; NADIRI, M.I.; ROLLER, L. Sickles RC (1993) Efficiency and productivity growth comparisons of European and US air carriers: a first look at the data. *Journal of Productivity Analysis*, 4, p. 115–125. 1993.

HAYAMI, Y. Sources of agricultural productivity gap among selected countries, *American Journal of Agricultural Economics* 51:564–575, 1969.

HAYAMI, Y.; RUTTAN, V.W. *Agricultural development: an international perspective*. John Hopkins University Press, Baltimore. 1971.

HAYAMI, Y.; RUTTAN, V.W. Agricultural productivity differences among countries, *American Economic Review*, 60, p. 895–911. 1970.

HELFAND, Steven M.; MAGALHÃES, Marcelo M.; RADA, Nicholas E. Brazil's Agricultural Total Factor Productivity Growth by Farm Size. Inter-American Development Bank, 2015.

IBGE. Instituto Brasileiro de Geografia e Estatística. Disponível em: <http://www.ibge.gov.br>>. Acesso em: set/dez 2015.

IPEADATA – Instituto de Pesquisa Econômica Aplicada. Disponível em: <www.ipeadata.gov.br>. Acesso em: set/dez 2015.

IRFFI, G. D. Semiárido brasileiro e políticas regionais: o caso do Fundo Constitucional de Financiamento do Nordeste (FNE). Relatório. IPEA, 2015.

KHAN, A. S.; CAMPOS, R. T. Effects of drought on agricultural sector of Northeast Brasil. *ICID*, 1992.

MARENGO, J. A. Vulnerabilidade, impactos e adaptação à mudança do clima no semiárido do Brasil. *Parcerias Estratégicas*. Brasília, n. 27, 2008.

MEEUSEN, W.; BROECK, V. D. Efficiency estimation from Cobb-Douglas production with composed error. *International Economics Review*, 32, p.715-723, 1977.

NASUTI, S.; EIRÓ, F.; LINDOSO, D. Os Desafios da Agricultura no Semiárido Brasileiro. *Sustentabilidade em Debate*, v. 4, n. 2, p. 276-298, 2013.

O'DONNELL C. J.; RAO, D.S.P.; BATTESE, G.E. Metafrontier frameworks for the study of firm-level efficiencies and technology ratios, *Empirical Economics*, 34, p. 231–255, 2008.

PITT, M. M.; LEE, L. F. The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry, *Journal of Development Economics*, 9, p.43-64, 1981.

Productivity Note, Ministry of Agriculture, Livestock, and Food Supply (Ministério da Agricultura, Pecuária e Abastecimento), Brasília, 2014.

ROCHA, R.; SOARES, R.R. Water Scarcity and Birth Outcomes in the Brazilian Semiarid. *Journal of Development Economics*, v. 112, p. 72 – 91, 2015.

SCHIMIDT, P.; SICKLES, R. Production Frontiers and Panel Data, *Journal of Business and Economic Statistics*, vl.2, n.4, p.367-374, 1984.

SIMÕES, A. F.; KLIGERMAN, D. C.; LA ROVERE, E. L.; MAROUN, M. R.; BARATA, M.; OBERMAIER, M. Enhancing adaptive capacity to climate change: the case of smallholder farmers in the Brazilian semi-arid region. *Environmental Science & Policy*, v.13, p.801-8, 2010.

VICENTE, J.R. Mudança tecnológica, eficiência, produtividade total de fatores na agricultura brasileira, 1970-95. *Economia Aplicada*, v. 8, n. 4, p. 729-760, 2004.

VIEIRA FILHO, J. E. R. Heterogeneidad estructural de la agricultura familiar en el Brasil. *Revista de la CEPAL (Impresa)*, v. 111, p. 103-121, 2013.

VIEIRA FILHO, J.E.R. Transformação histórica e padrões tecnológicos da agricultura brasileira. In: BUAINAIN, AM; ALVES, E.; SILVEIRA, JM; NAVARRO, Z. O mundo rural no Brasil do século, v. 21, p. 395-421. 2014

Wages and Net Fiscal Benefit in a Federation: Evaluating Intergovernmental Grants in Brazilian Municipalities

Ricardo Batista Politi (UFABC)*
Enlinson Mattos (EESP – FGV)

Abstract: In this article, we estimate the net fiscal benefit and the fiscal effect of unconditional grants on wages in metropolitan and non-metropolitan areas for 26 States from 2004-2009 period in Brazil. In particular, we explore the analytical framework proposed in Albouy (2012) to investigate the efficiency and equity effects of unconditional intergovernmental grants in Brazil. This framework can be understood as a positive exercise to comprehend whether grants mitigate the net fiscal benefit differences across localities and whether they promote equalization in certain criteria. Results suggest that grants distribution can be improved in both equity and efficiency. Our estimates of the net fiscal benefit indicate that the unconditional grant policy in Brazil is associated with inefficiency because higher paying areas in metropolitan regions are negative recipients of grants. More interestingly, we find that, contrary to conventional wisdom, grants distribution, although positively associated with certain minorities characteristics, is negatively related to our measures of inequality.

Keywords: Intergovernmental Transfers, Fiscal Federalism, Net Fiscal Benefit.

JEL Codes: H72, H77, R10.

**Os autores agradecem o apoio da FAPESP (processo nº 2012/23862-5).*

1. Introduction

The equalization aspects of intergovernmental grants have generated much debate in the public finance literature. In the search for stronger or weaker equalization, central governments encounter several issues regarding efficiency (see Buchanan 1950; 1952; Musgrave 1961; Boadway and Flatters 1982). A relevant debate in this literature is fiscal capacity equalization or the discussion over equalization based on source revenues compared with residence-based resources (Boadway 2004).

In this article, we estimate the net fiscal benefit and the fiscal effect of unconditional grants on wages in metropolitan and non-metropolitan areas for 26 States under the 2004-2009 period in Brazil. In that country, unconditional grants are the main source of revenues for local governments where much of the public provision is locally provided (Higgins et al. 2015). Moreover, local governments can engage in redistribution through the public provision of private goods such as health and education services at no cost to their citizens (Besley and Coate 1991; Boadway and Keen 2000)¹. This reinforces the importance of comprehending the redistributive and efficiency effects of grants on those municipalities². In particular, we explore the analytical framework proposed in Albouy (2012) to investigate the efficiency and equity effects of unconditional intergovernmental grants in Brazil.

In contrast to income inequality literature, in which most research approach direct transfers (cash transfers) and indirect tax in the income analysis (e.g., post-fiscal income), the current framework allows us to understand the fiscal impact of grants on location wages. In fact, if a worker lives in an area with net fiscal benefits above the national average, she has access to a publicly provided good at a lower cost; put differently, her net income rises more than the wage variation (refer to Albouy 2012, pp. 832). Previously, Beeson et. Al (2010) found that local fiscal policies are associated to wage distribution.

There is a broad body of literature that discusses the analytical basis for vertical equalization in a Federation. According to Buchanan (1950), the main aspect of equalization is to assure that individuals, mainly the poor, have access to public provisions separate from where they choose to live. In a broad sense, this type of redistribution typically means horizontal equalization; richer communities should provide resources to their poorer counterparts. As developed by Albouy (2012), the literature on fiscal federalism has discussed how it is possible to conciliate both efficiency and equalization objectives when the resources are available according to individual characteristics instead of community characteristics (e.g., poorer communities versus poorer individuals).³

In most equalization systems, unconditional grants are negatively associated with a measure of fiscal capacity. In fact, most developed countries use '*Representative Tax System*' (RTS) to pursue fiscal equalization. In this system, central governments attempt to estimate fiscal capacity and to promote equalization based on fiscal need. However, it is very difficult to determine fiscal need, and there is much discussion regarding whether revenue equalization should be a sole source or destination based. Despite this, few works provide straightforward models that consider local characteristics such as labor force composition and productivity, differences in local resources or costs, or income inequality to evaluate the equity and efficiency of grants (Boadway and Flatters 1982).⁴

An interesting aspect of Albouy's model is that, to address the net fiscal benefit, it considers potential earnings instead of nominal wages and incorporates a federal tax differential. According to this analytical framework, grants should mitigate the net fiscal benefit differences across localities in such a manner that, in equilibria, all households from different regions would encounter the same net fiscal benefit. The rationale for this belief is that grants should not incentivize workers who live in more productive labor markets to move to areas with lower productivity to obtain higher net fiscal benefits. Albouy (2012) maintains that living in a low paying area is different from being a lower paid worker. Thus, grants should be redistributed to areas with lower paid individuals.

¹ Besley and Coate (1991) provide a model to investigate first-best redistribution of publicly provided goods with imperfect information on the part of the government.

² Moreover, Boadway and Marchand (1995) find that overprovision of publicly provided goods could be welfare-enhancing.

³ However, because Federations distribute resources based on jurisdictions characteristics instead of household or individual characteristics, it is not possible to approach equity without producing inefficiency.

⁴ For a previous and comprehensive discussion of Fiscal Federalism in Brazil regarding equity and efficiency, see Shah (1991) pp. 92. Mendes et. Al (2008) provides a descriptive analysis in Portuguese.

Concerning equity, we expand the analysis because we estimate an inequality index (Gini Index) separately for each unit of observation, and we incorporate other inequality measures into the analysis, such as median wage and rank ordering of wage distribution. Moreover, we explore the fact that in Brazil, the minimum wage is an important bottom threshold for formal workers earnings to address the efficiency effects of grants. Finally, the distinction between fiscal benefits in metropolitan and non-metropolitan areas is an important issue in Brazil⁵ because there are nearly half the population with a huge demand for public services. Thus, in contrast to Albouy (2012), we examine metropolitan and non-metropolitan areas separately, and we include wage inequality indexes in the analysis. Regarding methodology, we obtain a suitable number of observations (61 areas in Brazil against 9 provinces in Canada) to develop panel data approach in which fixed idiosyncrasies of the data unit are removed.

Coefficient estimates of grants on location wage suggest that a worker that moves to increase her earnings will encounter a decrease in federal grants and a decrease in NFB (approximately 4% and 3%, respectively), decreasing her total gains. Although the inefficient effect of grants on location wage is low, the current policy of grants distribution works against formalization of the labor market. Regression coefficient of grants on the share of individuals with minimum wage suggest that if individuals move to metropolitan areas with a higher share of formal jobs, they will lose fiscal benefits. Regarding equity, grants do not reach minorities living in higher paying areas in contrast to minorities living in lower paying communities. Grants in Brazil are also not associated with local public provision costs. More importantly, our results indicate that the most populated metropolitan areas in Brazil are the most harmed in terms of the Net Fiscal Benefit.

To discuss equity and efficiency issues in grants distribution in Brazil, the paper is presented as follows. The next section provides the institutional background in Brazil, as well as our data. Section 3 details the Net Fiscal Benefit (NFB) specification and results. Finally, Section 4 provides the conclusion.

2. Institutional Framework in Brazil and Data

To approach grants and net fiscal benefits in Brazilian municipalities, we need to comprehend several issues, such as Fiscal Federalism, including sourced-based taxes as royalties, wage inequality and income tax. In this section, we provide a brief description on those topics.

In Brazil, municipalities have a large fiscal gap (difference between expenses and revenues). This phenomenon happens because although the central government has a great share of the total tax revenues, local jurisdictions have a small share of the total tax basis and have a great share of the public expenses. In fact, municipalities in Brazil provide the public services of elementary education, social care, less sophisticated health services (such as first aid) and infrastructure, including public transportation (excluding subway services), road maintenance and public leisure areas such as parks and local libraries.

In Brazil, as noted by Mendes et al. (2008), in contrast to other countries with a similar level of development, vertical grants correspond, on average, to approximately 65% of municipalities' total revenues (for comparison, in Mexico, the grants' share of local revenue resources are less than 30%). This large dependency on grants reinforces the relevance of evaluating the equity and efficiency effects on grants distribution from the central government to local jurisdictions.⁶

The most important budget resource to municipalities in Brazil are the block grants from central government labeled MPF (Fundo de Participação dos Municípios, in Portuguese). Resources from the MPF represent, on average, approximately 1.2% of the Gross Domestic Product (GDP) (Mendes et al. 2008). The main objective of the MPF is to address municipalities' fiscal gaps, redistributing resources from the Federal Tax on Consumption (primarily levied on sin goods such as cigarettes and alcohol, and on telecommunications and gas) and the Federal Income Tax (levied on individuals and corporations).

⁵ See also Shah (2012) for a discussion of the importance of grant resources on metropolitan areas.

⁶ There are two local taxes in Brazil: a property tax and a tax associated with local private provided services such as restaurants, shopping centers, and groceries, etc. On average, they represent less than 15% of municipalities' total revenues, being more important in large municipalities. We do not include these taxes in the analysis because they are residence based revenues. According to Albouy (2012), this type of tax should reflect the local cost of public provision. In fact, there is widespread agreement in the literature about equalization differences (or fiscal need) for only sourced based revenues (see Albouy pp. 827, 2012).

MPF grants distribution follows a complex formula. State capital municipalities (27 observations, including district capital Brasilia) receive 10% of total MPF resources according to population size and average inverse income. The remaining 90% of the resources is allocated to non-State capitals; however, municipalities with more or less than 142,000 inhabitants receive different treatment. Municipalities with less than 142,000 inhabitants obtain resources based exclusively on 17 different population cutoffs. In contrast, municipalities with more than 142,000 inhabitants obtain additional resources of 4% (from the 90% share) of the fund, which is also based on population size and average inverse income. On average, because of this intricate rule, less populated municipalities tend to obtain more grants per capita than more populated cities.

The differences between the average MPF per capita in large cities compared with small cities in Brazil has raised debate over different criteria for grants allocation in metropolitan versus non-metropolitan areas. Despite the expected differences between metropolitan and non-metropolitan areas' fiscal needs, there is minimal agreement in the literature on whether grants distribution should vary according community size or geographic location. Although the analytical basis for grants equalization is independent of community size, larger cities tend to have a higher number of local public services and a broader variety of services compared with small and rural communities (Reschovsky 2007; Slack 2007).

On the one hand, it is not efficient to provide large amounts of grants for inhabited and isolated geographic areas, which have high costs for public provision (Slack 2007). On the other hand, certain central governments could find it advantageous to provide incentives for migration from high to low inhabited areas when congestion costs in the former are higher than equalization costs in the latter (Dafflon 2007). This difference in public provision is primarily justified in the public finance literature by population density or more social-economic diversity across households.

The second most important grants for municipalities originate from the States (second tier of government). States return approximately 20% of the sales tax revenues to the municipalities where the tax was originally collected.⁷ This 'returning' grant is also a blocked grant and can be freely allocated; however, what makes this State grant unconventional is that sales tax in Brazil is both a source and destination based tax. In fact, sales tax revenues originate from residence consumption (destination based) and local production (source based). Although there is no information available regarding the share of destination and source-based revenues, we will incorporate sales tax into the analysis (refer to Figure 3 in Section 4).⁸

Furthermore, municipalities obtain resources from the use of territorial resources as royalties. The federal government collects approximately 10% of the oil production value. The criteria for royalties' distribution to States and municipalities follow geographic closeness; in fact, coastal distance to the oil installation is dependent on geographic coordinates and latitude. According to the National Agency of Oil (ANP, in Portuguese), in 2010, approximately 33% of royalties' total revenues were distributed to municipalities as follows: 29% to States and 28% were available to the central government. However, most of the oil extraction in Brazil is offshore, and it is highly concentrated in certain States. In particular, approximately 82% of the total oil production and 92% of offshore production occurs in the State of Rio de Janeiro (Monteiro and Ferraz 2010). Thus, certain municipalities in this State heavily benefit from this source-based grant.⁹

⁷ In fact, states redistribute 25% from total resources. However, 75% of these shares are source-based, and the remaining 25% follow another state criteria (could be equalization). See also Shah (1991), Section 2.61, for further details.

⁸ Additionally, municipalities obtain two categorical grants as follows: one for educational expenses and another for health expenses. Fundeb is the categorical grant for education in Brazil. It is mainly an equalization grant. Municipalities obtain resources according to their share of total enrollment in each State. The idea is to guarantee a minimum expense per student from elementary and secondary levels in any locality. However, the resources for this education categorical come mainly from municipalities (MPF grants) and States, with a small share (less than 10% on average) of central government resources. Each State manages its own Fundeb fund. Similarly, grants for the healthcare system (SUS, Sistema Único de Saúde, in Portuguese) aim for equalization. Resources come from a central fund that redistributes to States and municipalities based on levels of health services provided locally. Municipalities with more health services (e.g. prevention programs) or more sophisticated procedures (inpatient stay) receive more resources. There is also a minimum per capita transfer for all localities. Again, for municipalities, a portion of the health resources comes from MPF transfers. Because of these characteristics (unconditional transfer finance part of fund resources), we do not consider those categorical grants in the analysis. Moreover, we note that categorical grants represent, on average, a small per capita transfer compared with MPF grants.

⁹ See Monteiro and Ferraz (2010) and Serra (2005) for a comprehensive discussion of royalties in Brazil.

Regarding income, we briefly describe two subjects as follows: the federal income tax and income inequality. Income taxes in Brazil are collected by the central government. Nominal tax rates are progressive according to income ranges, and there are several tax deductible expenses, such as healthcare and education. Income ranges and deductible expenses values vary frequently. In general, most of the population is below the first tax range cutoff or is tax exempt. The next group is taxed at a 15% tax rate. The richest group encounters a 27.5% tax rate. The main criticism of the Brazilian income tax is that, although progressive, it does not capture income outside formal markets, and those incomes are not necessarily lower than national average.

Finally, Brazil is one of the most unequal countries in the world. Because of this scenario, several research studies have investigated the equality of opportunity and earning inequality in Brazil.¹⁰ Currently, the minimum wage is also significant to explain regional inequality.¹¹ Silveira Neto and Azzoni (2011) show the importance of real minimum wages increases in the north and northeast regions in Brazil, where most workers (not including the formal workforce) earn less than minimum wage, to the national income convergence. Additionally, certain authors note that minimum wage in Brazil can disincentive workers to seek jobs in the formal sector (Higgins et al 2015). Similarly, our current framework raises concerns regarding the inefficiency impact of grants in the formal sector: workers moving to regions with high formal jobs can obtain a lower NFB because of grants distribution.

We note that Brazilian states are divided into five geographic regions: North, Northeast, Center-West, Southeast and South. Figure A.6 in the Appendix provides a map of Brazil with States and their respective regions. Summarizing, the complexity of grants rules and their impact on wages reinforce the importance of using a methodology to evaluate the equity and efficiency aspect of the unconditional grants in Brazil. In Section 3, we detail our empirical approach for this issue.

2.1 Data

To approach the efficient and equalization aspects of resources distribution from the central government to local communities (municipalities), considering both metropolitan and non-metropolitan areas, we collect information from two different datasets. The Finbra dataset of the Secretary of National Treasure provides detailed information on community public finance such as tax revenues, grants revenues and population. These data include information on grant types such as royalties and data on sales tax in Brazil (ICMS), which is a mix of source and destination tax. Information from Finbra data is used to compound our measure of NFB.

Micro data regarding individual wages are available in PNAD (National Research on Households). PNAD is a sample dataset that contains information regarding socio-economic and demography characteristics of the individuals, such as gender, age, ethnic group, educational level (in years) and whether the individual has migrated. Regarding the labor market, PNAD data provides information regarding raw wages, industry sector and occupation, number of monthly working hours, and whether the worker is formal (an important issue in Brazil) and whether she is self-employed. We note that, to consider full time workers, we select individuals aged from 21 to 59.

In contrast to Finbra data, which provides data annually for each municipality, data on PNAD is available by States and regions (State capitals, metropolitan areas out of capitals, and non-metropolitan areas). It is possible to conciliate information from both data using IBGE (National Bureau of Economic and Geography) classification on States' regions.¹² Thus, information regarding approximately 5,400 municipalities is grouped into three types of regions as follows: metropolitan areas in States' capitals (hereafter MR1); metropolitan areas outside States' capitals (hereafter MR2); and non-metropolitan areas (hereafter R3),

¹⁰ For a discussion on inequality in Brazil and, mainly, for the impact of indirect taxes on inequality see Lustig et. Al (2014).

¹¹ As a matter of fact, minimum wage has an important influence on wages at urban labor markets, mainly for less educated-workers. For example, in the USA labor market, Easton (2006) concludes that minimum wage has a larger impact on the wages of workers with less than a high school degree than local industry mix.

¹² In fact, it is possible to compose 9 out of 13 Metropolitan areas (non State capitals) as used by IBGE. The remaining four areas are classified as "regions of development" and cannot be merged with PNAD data. Therefore, they are classified as R3 regions. Available at: << http://www.ibge.gov.br/home/estatistica/populacao/censo2010/sinopse/sinopse_tab_rm_zip.shtm>>.

including rural areas. Because of its particular characteristics, we exclude Brazil's capita district (Brasilia) from our sample. With this segmentation, our sample contains 61 different units of observations. This classification allows us to retain the analyses of potential differences between metropolitan and non-metropolitan areas and continue to compound a number of significant observations for a panel data approach.

To compose our panel, we use time data from 2004 to 2009.¹³ All monetary values are in Brazilian Reais (R\$) and were deflated to the year 2009 value based on the Consumer National Price Index (IPCA) available at IBGE. To estimate the federal tax differential (income tax) originating from location effects, we collect data on the nominal tax rates on wages in the Brazilian Income Revenue System.¹⁴ PNAD data allow us to estimate Federal tax differentials considering workers who are exempt from income taxes.

3. Evaluating Grants in Brazil

3.1 Specification of Location Wage and Net Fiscal Benefit (NFB)

Compared with Buchanan (1950), an important difference in Albouy's (2012) model is that although the first considers individuals with the same nominal wage as identical, independent of where they live, Albouy (2012) maintains that individuals living in the same locality and with the same wage should be treated as equals. In fact, to mitigate efficiency concerns when approaching equalization, one should consider potential earnings instead of nominal wages. Efficiency is an important issue because workers with higher productivity can be discouraged from searching for jobs in higher earning areas due to a federal tax differential or a negative NFB.

To approach potential earnings, Albouy's (2012) models depart from a standard *mincerian* wage function:

$$w_e^{ij} = X_e\beta + \mu^j + \varepsilon_e^{ij}$$

Where w_e^{ij} is the natural logarithm of wages of individual i in locality j , X_e is the vector of attributes for a worker of type e , and the term μ^j represents locality characteristics. Thus, the estimate of average potential earnings from a specific locality j follows function (3):

$$\bar{w}^j = \bar{X}^j\beta + \mu^j \quad (1)$$

Where the first term ($\bar{X}^j\beta$) is the predicted wage controlling for worker skills and the second term (μ^j) represents the locational effect on wages, independent of the local human capital level. In fact, to obtain wage fitted values, we multiply the estimated covariates coefficients by explanatory variable averages in each locality (see Appendix A.1 for details). We should note that the total fitted wages and their two components, location and composition effects are normalized to obtain a national average of zero. We use the same procedure (normalization) for all components of NFB as in Albouy (2012).

TABLE 1 HERE

Table 1 provides per capita wage differentials compared with the national average as weighted by population. We use three main areas of analysis, as described previously in Section 2: metropolitan areas in States' capitals (hereafter MR1); metropolitan areas outside States' capitals (hereafter MR2) and non-metropolitan areas (hereafter R3). Wage differentials note that the highest nominal wages in Brazil are in the metropolitan areas (MR1 and MR2) of São Paulo State (SP), in the States of the South region such as Paraná (PR) and Rio Grande do Sul (RS), and in the MR1 areas of Rio de Janeiro (RJ) and Santa Catarina (SC).

¹³ We note that at the time of data collection, PNAD data were available only until the year 2011. We do not use data for the year 2010 because in census years, household surveys use different population weights.

¹⁴ We collect the respective nominal tax rates and deduction amounts at Brazilian Internal Revenue Service. Available at: <http://www.receita.fazenda.gov.br/Pagamentos/PgtoAtraso/tbcalcir.htm#2004>. We estimate effective income tax by regions using wages microdata available at PNAD.

Our estimates of location wage in accordance with Function (3) appear robust. According to the results in Table 1, column 2, localities in metropolitan regions with well-known higher living costs present a higher location wage; this is the case in the MR1 areas of de São Paulo, Paraná, Santa Catarina, Minas Gerais (MG), Rio de Janeiro e Goiás (GO).¹⁵ Although they have a lower differential, the same is true for MR2 communities in the Southeast and South regions. Second, communities well known for poor amenities that are also far from the wealthiest areas from the Southeast region also have higher location wage estimates. That is the case for the MR1 areas of Mato Grosso (MT) and Acre (AC) e Rondônia (RO). The remaining localities, primarily in R3 areas, present a negative differential of location wage compared with the average; this appears reasonable with the idea of a lower cost of living outside metropolitan areas, or in certain cases, a higher quality of living, such as lower crime rates.¹⁶ The MR1 areas in Mato Grosso, Paraná, Rio de Janeiro, and São Paulo also encounter difficult commuting; this tends to be compensated for by higher wages.¹⁷

In contrast, the wage composition effect is associated with worker productivity (e.g., workers sort on localities based on their skill levels, Albouy, 2012 pp. 831). According to our estimates (function 1), a positive composition wage differential is found in localities with higher average years of schooling (see Table A.3 in the Appendix). Thus, a higher wage differential due to higher productivity in the workforce is found in all areas in States in the South region and in São Paulo and Rio de Janeiro. In particular, the differential is higher in the metropolitan areas (MR1 and MR2) than in other areas (R3). Additionally, as expected, localities in regions well known for lower educational levels (and lower worker productivity) present negative differentials on composition wages, as do the States in the Northeast region.

FIGURE 1 HERE

According to our estimates, location and composition effects have similar shares for total wage composition. Considering all areas (MR1, MR2 and R3), location effects represent approximately 47% of total wage earnings, and composition effects represents the remaining 53%. Additionally, we run a regression of composition wage on total raw-wage differences and obtain an estimated coefficient of 0.20 (with robust standard error of 0.019). This positive relation suggests that, in Brazil (in contrast to Canada, Albouy 2012, pp. 831), workers sort significantly across communities: higher paying areas attract workers with higher potential wages and more years of schooling. According to Figure 1, R3 communities are typically lower paying areas.

Interestingly, Figure 2 depicts a regression of the predicted wage differences of minorities (we solely consider the estimated coefficients for independent variables that describe minorities, such as women, immigrants, and individuals with informal jobs, see Appendix A.1) on raw wages and find that higher-paying areas also attract minorities. It is noteworthy to note that five of nine MR2 areas are on the superior right side of Figure 2. This finding suggests that those areas, which are a negative recipient of grants, also attract workers with lower paying characteristics who consider the earnings potential. Thus, metropolitan regions from the South and Southeast (MR1 and MR2) typically pay more to minorities than R3 areas. Considering equity, it is important to comprehend whether areas with more (less) grants than the national average host more (less) individuals with lower paying characteristics.

FIGURE 2 HERE

Regarding efficiency, federal tax rates on wages can influence individuals to move to areas with lower nominal wages. Thus, to avoid distortion of labor force allocations, federal grants should be used to offset this effect by generating fiscal benefits to highly taxed communities. Albouy (2012) develops an analytical set-up in which grants balance the effect of the three main fiscal components as follows: differences in federal fiscal

¹⁵ For example, Rio de Janeiro and São Paulo were ranked as the 19th and 26th most expensive cities in the world in 2010. <http://www.citymayors.com/statistics/expensive-cities-world.html>.

¹⁶ See Sachsida et Al. (2007) for an analysis of crime rates and urbanization in Brazil.

¹⁷ See Timothy and Wheaton (2001) for a discussion of commuting time and wages variation. For data on commuting in Brazilian metropolitan areas see

<<http://www.ibge.gov.br/home/presidencia/noticias/imprensa/ppts/0000000841420412201242262753970.xls>>

burden, ‘fiscal residuum’, and public good externality.¹⁸ The first component approaches the federal tax differential that originated from the wage variation.

The second fiscal component evaluates, compared with the national average, source based revenue differentials and local public expenses. Regarding efficiency, intergovernmental grants should compensate for source-based revenues as royalties. The main idea is that community-based taxes should reflect local costs in a manner that residence taxes work as user fees. Finally, the third fiscal component in Albouy’s model is the public good externality. This term reflects the effect of a new residents in the community. Based on the idea of non-rivalness of the public good, the new migrant increases local expenditures by an amount larger than her consumption. However, the literature on externality estimates that its effect is very small (Bergstrom and Goodman 1973).

TABLE 2 HERE

In order to obtain the Net Fiscal Benefit (NFB) for the habitants of each locality, we need first to estimate each component separately. Table 2 provides the per capita differential of the three fiscal components relative to the national average. Per capita values are weighted by population regarding the three main regions of analysis (MR1, MR2 and R3). Table 2 indicates that MR2 is the lowest net recipients of unconditional grants with a negative differential of R\$ 43 on average per capita, followed by MR1 areas that receive, on average, R\$ 10 less than average. In contrast, R3 regions are net positive recipients of unconditional grants with a positive differential of R\$ 103 compared with the national average. Despite those differences, there are also large differences in per capita grants within groups. For example, in area RM1, certain States in the North Region, such as Acre (AC), Roraima (RR) and Tocantins (TO), obtain much more grant resources than the national average, with a positive surplus of approximately R\$ 200 to R\$ 500 per capita. Other States from the South and Southeast, such as Rio Grande do Sul (RS) and Rio de Janeiro (RJ) e São Paulo (SP), receive less than the national average in both metropolitan areas, MR1 and MR2. In contrast, localities in R3 areas obtain more grants per capita than the average in all States, with the exception of Rio de Janeiro e Roraima.

We should note that there are areas in all regions of Brazil (e.g., North, Northeast, Center-West, Southeast, and South) that are positive recipients of unconditional grants compared with the national average. This is the case for Goiás (GO) and Mato Grosso do Sul (MS) in the Center-West region, Tocantins (TO) in the North region, Rio Grande do Sul (RS) and Paraná (PR) in the South region, Minas Gerais (MG) in the Southeast region and Alagoas (AL), Paraíba (PB), and Piauí (PI) e Rio Grande do Norte (RN) in the Northeast region. Regarding those five Brazilian regions, the North region is the highest net recipient as a whole, and the Northeast is the highest net recipient outside metropolitan areas (R3). Considering the metropolitan regions (MR1 and MR2) solely, the Southeast region is the lowest net recipient of federal block grants.

Table 2 provides the differential of the two remaining components of NFB. Regarding a source-based tax differential, column 2 provides localities with a differential on royalties associated with oil extraction. As previously discussed in Sections 2, sourced-based resources are inefficient because they affect local tax prices and are not associated with any fiscal need. In Brazil, the resources are heavily distributed to municipalities in Rio de Janeiro (refer to Section 2), primarily to non-metropolitan (R3) areas. In fact, the royalties differential is very large compared with the national average as outside Rio de Janeiro, five other areas of 58 (MR1, MR2 and R3) obtain a positive royalties differential.

As expected, metropolitan regions with higher wages and a higher share of formal jobs collect more in federal income tax compared with the national average. In spite of large wage differentials, the income tax differential is not that high; it is higher on the metropolitan areas of Rio de Janeiro and São Paulo (MR1) and lower in R3 areas in the North and Northeast States. Because income tax follows a Federal tax rate, the differential compared with the national average reflects how much the inhabitants of a specific area collect more or less in taxes regarding other areas as a consequence of local labor market characteristics as a share of formal jobs and a share of wages below the first income tax rate threshold.

Importantly, we note that areas outside metropolitan areas have a greater share of workers with earnings below the minimum wage. Column 4 in Table 2 shows that although the metropolitan areas of MR1

¹⁸ See also Appendix A.2 for a brief description of Albouy (2012). In contrast, we disregard potential public provision externality and congestion effects, since they are probably small in most areas. We should note that Albouy’s (2012) analysis also includes local cost of living and amenities. However, we do not have this information available in Brazil for our 61 units of observation.

and MR2 have approximately 20% of the workforce earning less than minimum wage, in R3 areas, that share rises to 40%. We also note that although the Gini Index (Column 3 in Table 2) is high in all localities, it is increased in MR1 areas. This finding likely reflects the fact that in metropolitan areas, the range between the highest wages and the lowest wages is higher, although the average wage is not.

We use these three fiscal components to compute the per capita differential or NFB weighted by population as follows: grants, source-based tax and federal income tax. Estimates of the NFB differential are most influenced by grants; therefore, on average, RM1 localities are the most damaged, with a negative gap of R\$ 45 in NFB per capita compared with the national average. Similarly, RM2 regions present a negative NFB gap, which is fairly small at approximately R\$ 29 per capita. In contrast, R3 areas are the localities, which, on average, benefit most from the current federal grants policy; they present a positive NFB differential of approximately R\$ 168 per capita. This scenario is intensified by the income tax differential, because most federal taxes are collected in metropolitan areas (MR1 and MR2) where the nominal wage is higher compared with the national average. Additionally, royalty revenues also accentuate this disequilibria because four R3 areas obtain more than R\$ 50 per capita compared with the national average (Sergipe [SE], Rio Grande do Norte [RN] and Espírito Santo [ES] and mainly Rio de Janeiro).

Although incomplete, this analytical framework allow us to evaluate grants distribution in terms of equity and efficiency, and more importantly, to evaluate which trade-offs are involved in grants distribution and how to balance the trade-offs that grants generate (refer also to Boadway and Flatters 1982). In the next section, we present our estimate for the Brazilian case.

3.2 Evaluating Equity and Efficiency

Using our estimates of wages, inequality indexes and data on grants differentials, we analyze NFB in accordance with the analytical framework discussed in Section 3.1. Our main results are presented in Table 3. We conduct regressions of unconditional grants and NFB on variables that are associated to efficiency (location wage and minimum wage), to inequality (Gini index, wage ranking distribution and median wage) and to public cost provisions (density and public sector wages). Regarding efficiency, we would expect that differences in the federal income tax would be offset by federal grants, or that higher-paying areas are not punished by negative NFB. Additionally, we calculate the difference between the mean and median wages, and we rank the ordering of individual wages and obtain the average in each locality compared with the national distribution. If those two indicators are associated with equity, one should expect that municipalities with a higher wage ranking and larger difference between the mean and median wage exhibit negative and positive relations to grants, respectively.

TABLE 3 HERE

As discussed in Albouy (2012), equity criteria for grants distribution considers that communities with lower potential income, because of the composition of the workforce and the mix of economic activities, should obtain more fiscal resources. Thus, location wages describe the effect of the community on wage determination, discarding the effect of human capital. The idea is that the location effect reflects the local cost of living or the amenities and the quality of life. One should expect that firms in communities with higher living costs or lower amenities would pay a higher wage to maintain its workers.

Coefficient estimates of grants regarding location wage suggest that a worker that moves to increase her earnings in one R\$ will encounter a 0.04 cent decrease in federal grants and a 0.03 decrease in NFB, such that her gains will rise by slightly less than the total gains, by approximately 0.97 cents. This means that there is a disincentive effect to move workers to higher paying areas because they will obtain less in fiscal benefits. Moreover, because unformal jobs are an important issue in labor market in Brazil, an important boundary on market place is the minimum wage. Regression coefficients of grants regarding individuals with minimum wages suggest that a community that creates formal sector jobs will obtain less grants and less NFB. In fact, this result reinforces the idea that, if individuals move from R3 regions (with a lower share of informal jobs, see Table 1) to metropolitan areas (MR1 and MR2) with a higher share of formal jobs, they will lose fiscal benefits. In other words, municipalities with higher than average location wages will have less resources per

capita for public provision than low-paying areas. Although this difference is significant in most regressions (columns 1 to 3), the magnitude is much lower than those found in Albouy (2012).¹⁹

The results on predicted earnings for minorities (as women, immigrants, non-white and informal workers) suggest a positive correlation between grants and potential earnings in this group. Although this suggests equality in grants distribution, the higher than average grant benefit for this group is offset by the negative NFB. Overall, this result supports the findings in Figure 2 as follows: if a minority group moves from a low-paying area to a high-paying area, they will encounter a negative NFB; in particular, they will lose approximately 0.04 cents in fiscal benefits for each additional one Real (R\$) in formal wages. In this case, the difference between grant and NFB arises from the federal income tax.

Regressions of grants and NFB on inequality indicators present mixed results. The negative coefficient (but not significant) in Column 4 (Table 3) suggests that grants have equity: if workers move to communities with higher (lower) rank ordering for average wage, they obtain fewer (more) grants. However, a higher inequality measure as the difference from mean wage to median wage is negatively associated with grants distribution. The same results are found for the Gini Index: if the worker moves to a community with higher inequality, she receives fewer intergovernmental grants.

Coefficients of NFB regressions on inequality index are similar to the grant results: they are negatively related (although not significant) to the mean to median difference and close to zero in the Gini Index. More surprisingly, because of royalties' effects on NFB, the coefficient on rank ordering on average wage is significant and positive: workers moving to areas with a higher ordinal position in wages obtain more fiscal benefits, which again moves in the opposite direction of equity.²⁰ The positive association between grants and inequality indicators is surprising because, due to the inverse income criteria on grants distribution, there is widespread belief that grants are positively related to poorer areas (Mendes et al. 2008). Thus, grants distribution in Brazil does not appear to be strongly related to inequality. In fact, this distribution works against the minorities living in higher-paying areas.

Finally, we investigate the association between grants and local public costs to determine whether grants benefit communities with higher public provision costs. The coefficient regression on density suggests that localities with higher density obtain less grants per capita. However, this result reflects much of the situation in which highly populated metropolitan areas (MR1 and MR2) are negative grants recipients. In fact, if we run a separate regression with R3 areas solely, the coefficient estimate is highly insignificant (p-value of 0.82), although it remains negative.

More interestingly, we run a regression of the sum of wages on public administration and health and education workers against total wages. Because a large share of public provision is on health and education services, this variable should provide a suitable comparison between local costs and the national labor cost differential. The coefficient estimate indicates a positive association between the average wage per capita and the average wage on public sector type activities (see also Figure A.5 in the Appendix). This suggests that labor costs in the public sector are associated with private sector wages. The results on density and public sector wages do not provide us with evidence that grants are directed to most needed communities in terms of costs.

FIGURE 3 HERE

In Brazil, there is an additional debate regarding whether grants should equalize part of the revenue differences in State sales tax (see also Shah 1991, Section 3.3; the debate is still present). This discussion also address the issue of residence-based versus source-based tax, because an idiosyncrasy of sales tax in Brazil is that it follows both criteria (refer to the discussion in Section 2). Although data on the sourced or residence-

¹⁹ One possible explanation for the great difference in coefficient magnitudes is that in addition to the differences between Brazil and the Canada Federation System, we run panel data models different from Albouy (2012), who runs a cross-section analysis. Pooling regressions coefficients (reported in Table A.4 in the Appendix) are larger in all regressions and have a different sign for two regressions (for minimum wage and minority wage). We argue that this difference between fixed effects and pooling functions reflects unobserved characteristics that vary across localities, such as institutional and business environments, or local residence tax rates levels that could bias the estimated coefficients.

²⁰ In fact, if we exclude the R3 area of Rio de Janeiro, which is a great outlier on royalty per capita distribution, we obtain a negative and not significant coefficient for the Gini Index variable.

based share of the tax are not available, we assume that sales taxes have an important relationship with fiscal capacity. We argue that the higher the sales tax per capita, the higher the fiscal capacity. Figure 3 plots the scatter of unconditional grants per capita (vertical axis) against the sum of the sales tax plus royalties and the discounted federal income tax. In this plot, the distance between each community marker (dot) and the solid dark line represents the NFB. More importantly, the distance provides the grants amount that would be needed to offset any fiscal disequilibrium, or to make measurable NFB zero (Albouy 2012).

Thus, communities to the right and superior quadrant of the solid line have a positive NFB; furthermore, communities in the left and inferior quadrant of the solid line have a negative NFB. The slope of the dashed line brings the coefficient of the per capita grant against the set of the other previously described fiscal components (royalties, federal income tax, sales tax). The slightly negative slope of the dashed line suggests that grants distribution does not balance NFB between communities, nor does it address fiscal capacity. More interestingly, we note that approximately two thirds of the localities are below the solid line, which suggests that the actual rule of grants distribution benefits only a few localities.

Overall, we note that the markers more distant from the solid line are primarily from the MR1 and MR2 areas. Close inspection of Figure 3 suggests that RM1 of Rio de Janeiro should obtain approximately R\$ 300 more per capita to offset a negative NFB. Several other MR1 localities from all regions, such as Santa Catarina (South), Goiás and Mato Grosso do Sul (Center West), Ceará (Northeast) and São Paulo (Southeast) exhibit a high deficit in NBF. Most damaged MR2 localities are in the South (Paraná and Rio Grande do Sul) and Northeast (Ceará, Pernambuco e Bahia). Approximately 44 million inhabitants, or approximately 23.5% of the total population, live in these MR1 and MR2 regions.

With respect to efficiency, a positive or negative NFB affects worker migration because advantageous fiscal communities attract habitants in spite of labor market characteristics (e.g., lower versus higher paying areas). Thus, according to Albouy (2012), workers living in higher wage areas are punished for living in localities with higher productivity. Moreover, grants should be used as a corrective policy to offset the higher tax burdens of workers who live in areas with higher productivity. Regarding equity, fiscal components and NFB do not appear to be associated with fiscal capacity or fiscal need in Brazilian municipalities, which works in the opposite direction of equality. We also find a weak association between the local cost of public provisions and grants.

To summarize, evaluating grants in Brazil based on both equity and efficiency criteria suggests that although the inefficient effect of grants on location wage is low, the current policy of grants distribution works against the formalization of the labor market (as in Higgins et al 2015). For equity, grants do not reach minorities living in higher paying areas, in contrast to minorities living in lower paying communities. More surprisingly, grants distribution is negatively associated with inequality indices on wages, which reinforce the well-known idea that to address inequality, direct cash transfers are more suitable than intergovernmental transfers.

4. Conclusion

This paper approaches equity and efficiency in grants distribution in Brazil. Our analysis utilizes the analytical framework proposed by Albouy (2012). This approach is the first to consider that central government grants should mitigate potential misallocation effects associated with higher federal tax burdens on higher wages. The main implication is that individuals living in higher paying areas because of higher productivity should not be punished with lower net fiscal benefits compared with the national average. In other words, federal grants should offset allocation distortions that stimulate higher paying individuals to migrate lower-paying areas because of lower productivity.

Additionally, this analytical reference allows us to address equity. We investigate how grants are associated to workers with lower-paying characteristics, such as minorities, and more general inequality measures across localities, including the Gini Index, ordinal ranking of average wages and mean-to-median wages differences. Thus, in contrast to Albouy (2012), in addition to inequalities indexes, we group our sample of approximately 5.4 thousand municipalities across 26 Brazilian States into three main areas as follows: metropolitan areas in States capitals, metropolitan areas outside States capitals and non-metropolitan areas. It is important to comprehend grants effects on large urbanized areas in contrast to small areas in Brazil because

there is high demand for public services and greater income inequality. Regarding methodology, this empirical strategy led us to 61 units of observation and allowed us to control for fixed effects.

Our estimates of the net fiscal benefit indicate that the unconditional grant policy in Brazil is associated with inefficiency because higher paying areas in metropolitan regions are negative recipients of grants. More interestingly, we find that contrary to conventional wisdom, grants distribution, although positively associated with certain minority characteristics, is negatively related to our measures of inequality. Results suggest that grants distribution in Brazil can be improved in terms of both equity and efficiency. This type of analysis would benefit from additional data on amenities and the local cost of living.

5. References

- Albouy, D. 2012. "Evaluating the Efficiency and Equity of Federal Fiscal Equalization", *Journal of Public Economics*, 96(9-10): 824-839.
- Beeson, Patricia, Lara Shore-Sheppard, and Tara Watson. 2010. "Local fiscal policies and urban wage structures." *Public Finance Review*, 38(5) 540-584.
- Besley, T. and Coate, S. 1991. "Public provision of private goods and the redistribution of income", *American Economic Review*, September, 81, 979-993.
- Bergstrom, T.C., Goodman, R.P. 1973. "Private demands for public goods", *American Economic Review*, 63, 280-296.
- Boadway, R.W. 2004. "The Theory and Practice of Fiscal Equalization", *CESifo Economic Studies*, 50 (1): 211-54.
- Boadway, Robin; Flatters, Frank. 1982. "Efficiency and equalization payments in a federal system of government: a synthesis and extension of recent results", *Canadian Journal of Economics* 15, 613-633.
- Boadway R. and Keen, M. 2000. "Redistribution", in A.B. Atkinson & F. Bourguignon, *Handbook of Income Distribution*, vol.1, pp.677-789.
- Boadway, R. and Marchand, M. 1995. "The Use of Public Expenditures for Redistributive Purposes", *Oxford Economic Papers*, 47(1), 45-59.
- Buchanan, James. 1950. "Federalism and fiscal equity", *American Economic Review* 40, 583-599 (Sep.).
- Buchanan, James. 1952. "Federal grants and resource allocation", *Journal of Political Economy* 60, 208-217 (Jun.).
- Dafflon, B. 2007. Fiscal Capacity Equalization in Horizontal Fiscal Equalization Programs. In: Boadway, R. e Shah, Anwar. *Intergovernmental fiscal transfers: principles and practice*. Ed. The International Bank for Reconstruction and Development / The World Bank (Washington, DC).
- Easton, Todd. 2006. "Metropolitan Wage Levels of Less-Educated Workers: 1986-1999." *Industrial Relations*, 45 (2) (April): 119-46.
- Higgins, S., Lustig, N., Ruble, W. and Smeeding, T. 2015. "Comparing the Incidence of Taxes and Social Spending in Brazil and the United States." *Review of Income and Wealth*, DOI: 10.1111/roiw.12201.
- Lustig, N., C. Pessino, and J. Scott. 2014. "The Impact of Taxes and Social Spending on Inequality and Poverty in Argentina, Bolivia, Brazil, Mexico, Peru and Uruguay," *Public Finance Review*, 42, 287-303.
- Mendes, Marcos; Miranda, Rogério B. and Blanco, Fernando. 2008. "Transferências Intergovernamentais no Brasil: Diagnóstico e Proposta de Reforma", *Texto para Discussão 40*. Consultoria de Orçamento do Senado Federal.
- Monteiro, Joana, and Ferraz, Claudio. 2009. "Resource Booms and Voracious Politicians: The Effects of Oil Shocks on Patronage, Rent-Seeking, and Elections", unpublished, PUC-Rio.
- Musgrave, Richard A. 1961. "Approaches to a fiscal theory of political federalism". *Public Finances: Needs, Sources, and Utilization*. National Bureau of Economic Research, New York.
- Reschovsky, A. 2007. "Compensating Local Governments for Differences in Expenditure Needs in a Horizontal Fiscal Equalization Program." In: Boadway, R. e Shah, Anwar. *Intergovernmental fiscal transfers: principles and practice*. Ed. The International Bank for Reconstruction and Development / The World Bank (Washington, DC).
- Sachsida A., M.J.C., Mendonça, P.R.A., Loureiro, M.B.S Gutierrez. 2007. "Inequality and criminality revisited: further evidence from Brazil", *Empirical Economics*, 39 (2010) 93-109.

Serra, R.V. 2005. "Concentração espacial das rendas petrolíferas e sobre financiamento das esferas de governo locais: evidências e sugestões para correção de rumo", XXXIII Encontro Nacional da ANPEC, Natal (RN). Available at: <http://www.royaltiesdopetroleo.ucamcampos.br/index.php?cod=4>

Shah A. 1991. "The new fiscal federalism in Brazil", DP 124, World Bank, Washington, DC.

Shah, A. 2012. "Grant financing of metropolitan areas: A review of principles and worldwide practices", Policy Research Working Paper 6002. Washington, D.C.: World Bank.

Silveira Neto, R. and Azzoni, C. 2011. "Non-Spatial Government Policies and Regional Income Inequality in Brazil", *Regional Studies*, Vol (45): 4 pp. 453-461.

Slack, E. 2007. "Grants to Large Cities and Metropolitan Areas". In: Boadway, R. e Shah, Anwar. *Intergovernmental fiscal transfers: principles and practice*. Ed. The International Bank for Reconstruction and Development / The World Bank (Washington, DC).

Thimoty, D. and Wheaton, W. 2001. "Intra-urban wage variation, employment location and commuting times", *Journal of Urban Economics*, 50(2), 338-66.

TABLE 1

Per capita wage differences across localities (total, composition and location effects) and inequality indicators

	<i>Differential Total Wage (level)</i>			<i>Differential Composition Effect</i>			<i>Differential Location Effect</i>			<i>Gini Index</i>			<i>% workers with earnings below minimum wage</i>		
	Metropolitan area		Non metropolitan area (R3)	Metropolitan area		Non metropolitan area (R3)	Metropolitan area		Non metropolitan area (R3)	Metropolitan area		Non metropolitan area (R3)	Metropolitan area		Non metropolitan area (R3)
	Capital (MR1)	Non cap. (MR2)		Capital (MR1)	Non cap. (MR2)		Capital (MR1)	Non cap. (MR2)		Capital (MR1)	Non cap. (MR2)		Capital (MR1)	Non cap. (MR2)	
<i>State</i>															
AC	310		-396	0.08		-0.38	0.29		-0.15	0.52		0.44	0.18		0.52
AL	116		-508	0.05		-0.42	0.14		-0.38	0.56		0.43	0.25		0.57
AM	146		-287	0.06		-0.21	0.16		-0.12	0.43		0.42	0.10		0.36
AP	149		-172	0.09		-0.16	0.14		0.01	0.44		0.38	0.13		0.26
BA	82	-142	-480	0.03	-0.05	-0.41	0.13	-0.06	-0.31	0.53	0.48	0.45	0.22	0.29	0.56
CE	-60	-121	-558	0.02	-0.06	-0.38	-0.02	-0.02	-0.56	0.50	0.53	0.50	0.25	0.33	0.68
ES	200		-206	0.13		-0.05	0.14		-0.15	0.49		0.44	0.14		0.31
GO	254		-91	0.13		-0.05	0.19		0.01	0.49		0.46	0.11		0.25
MA	70		-434	0.08		-0.36	0.06		-0.26	0.53		0.51	0.26		0.57
MG	268	164	-164	0.15	0.18	-0.03	0.19	0.06	-0.10	0.50	0.48	0.46	0.13	0.14	0.29
MS	283		-7	0.16		0.00	0.19		0.06	0.51		0.48	0.16		0.24
MT	405		-7	0.16		-0.06	0.29		0.12	0.50		0.45	0.11		0.25
PA	-27	-46	-288	0.02	-0.04	-0.28	0.02	0.05	-0.05	0.49	0.49	0.43	0.24	0.30	0.39
PB	41		-408	0.08		-0.29	0.04		-0.27	0.54		0.50	0.25		0.54
PE	1	-165	-508	0.10	-0.05	-0.35	-0.03	-0.08	-0.45	0.52	0.48	0.44	0.23	0.33	0.57
PI	73		-575	0.04		-0.50	0.11		-0.49	0.57		0.58	0.30		0.72
PR	476	492	-37	0.28	0.28	0.09	0.23	0.24	-0.06	0.48	0.51	0.46	0.10	0.12	0.26
RJ	356	145	-51	0.24	0.18	0.11	0.18	0.05	-0.09	0.50	0.47	0.42	0.10	0.13	0.16
RN	129		-451	0.14		-0.26	0.07		-0.38	0.53		0.44	0.20		0.52
RO	224		-207	0.09		-0.17	0.22		-0.03	0.49		0.46	0.16		0.38
RR	101		-360	0.07		-0.24	0.11		-0.24	0.48		0.48	0.18		0.50
RS	395	327	-90	0.29	0.26	0.12	0.15	0.13	-0.15	0.49	0.50	0.44	0.11	0.18	0.29
SC	586		171	0.36		0.21	0.23		0.05	0.46		0.43	0.07		0.17
SE	103		-421	0.07		-0.31	0.11		-0.27	0.52		0.44	0.21		0.50
SP	528	573	159	0.25	0.32	0.22	0.30	0.26	0.02	0.49	0.49	0.43	0.09	0.10	0.12
TO	239	-301		0.15		-0.16	0.16		-0.19	0.50		0.48	0.15		0.44
Total										0.50	0.49	0.45	0.17	0.21	0.40

Obs: Number of observations 61, with 26 metropolitan areas in States' capitals, 9 metropolitan areas outside States' capitals and 26 non metropolitan areas. Averaged values from 2004 to 2009 measured in 2009 Brazilian Reals (R\$). Composition and Location effects are expressed in logarithmic terms.

TABLE 2
Differential in per capita fiscal components - relative to national average in 2009 Brazilian Reais (R\$)

	<i>Population (year 2009)</i>			<i>Federal tax (1)</i>			<i>Royalties (2)</i>			<i>Grants (3)</i>			<i>Net Fiscal Benefit (1+2+3)</i>		
	Metropolitan area		Non metropolitan area (R3)	Metropolitan area		Non metropolitan area (R3)	Metropolitan area		Non metropolitan area (R3)	Metropolitan area		Non metropolitan area (R3)	Metropolitan area		Non metropolitan area (R3)
	Capital (MR1)	Non cap. (MR2)		Capital (MR1)	Non cap. (MR2)		Capital (MR1)	Non cap. (MR2)		Capital (MR1)	Non cap. (MR2)		Capital (MR1)	Non cap. (MR2)	
<i>State</i>															
AC	370,809		320,323	4		-10	-27		-23	194		24	53		109
AL	1,160,393	595,148	1,362,625	4		-10	-14		-5	34		145	9		261
AM	2,042,185		1,331,300	1		-8	-10		25	-130		68	-183		163
AP	463,704		155,103	-7		-11	-25		-24	23		38	-76		59
BA	3,866,004	367,061	10,339,889	8	-5	-9	-7	-20	-16	-101	4	93	-177	27	118
CE	3,655,259	560,325	4,332,225	-2	-1	-10	-21	-26	-19	-36	21	109	-71	83	151
ES	1,686,045		1,801,154	10		-7	-7		51	-113		81	-139		214
GO	3,198,918		2,723,702	2		-7	-27		-7	-73		173	-78		117
MA	1,439,021	334,899	4,521,769	5		-9	-26		-25	3		74	-33		123
MG	5,750,522	615,117	13,637,847	13	4	-5	-18	-24	-10	-96	61	122	-97	14	90
MS	755,107		1,605,391	4		-7	-29		-12	-135		154	-30		204
MT	823,831		2,177,861	3		-5	-25		-20	-99		101	-99		117
PA	2,161,191	479,714	4,484,599	-2	-3	-8	-27	-20	11	-80	-60	21	-154	-101	48
PB	1,156,745	687,545	1,905,021	-3		-9	-23		-24	73		207	45		241
PE	3,768,902	360,283	4,677,963	5	-7	-10	-22	-27	-21	-82	-22	97	-182	-79	95
PI	984,931		2,135,862	0		-10	-26		-23	29		180	65		224
PR	3,307,945	1,359,498	6,014,623	21	8	-5	-26	-27	-14	-99	-58	141	-143	-74	89
RJ	11,863,799	772,013	3,347,721	22	9	-3	8	259	770	-185	-91	-34	-218	193	763
RN	1,322,984		1,798,173	3		-10	-7		52	-30		231	-18		347
RO	382,829		1,121,099	-2		-9	-25		-24	63		15	-23		8
RR	278,843		142,656	-8		-11	-20		-26	515		-25	479		39
RS	4,064,186	1,408,160	5,441,782	16	11	-6	-25	-19	-15	-105	-127	190	-123	-162	139
SC	977,370	3,878,800	1,262,573	16		0	-27		-17	-32		66	-149		74
SE	794,475		1,225,204	3		-10	16		55	-13		144	34		316
SP	19,777,129	4,439,290	17,154,385	37	28	2	-28	-27	-18	-192	-114	18	-292	-160	-7
TO	309,353		982,698	-2		-10	-13		-18	409		248	439		272
Total	76,362,480	15,857,853	96,003,548	6	5	-8	-18	8	23	-10	-43	103	-45	-29	168

Obs: Number of observations 61, with 26 metropolitan areas in States' capitals, 9 metropolitan areas outside States' capitals and 26 non metropolitan areas. Averaged values from 2004 to 2009, weighted by population.

TABLE 3 - Relating grants and net fiscal benefit to efficiency and equity criteria

[illegible]

Obs: Number of observations is 305. Panel regressions using 61 localities from year 2005 to 2009, weighted by population. In parentheses are robust t-statistics.
* Significant at 10%; **significant at 5%; ***significant at 1%.

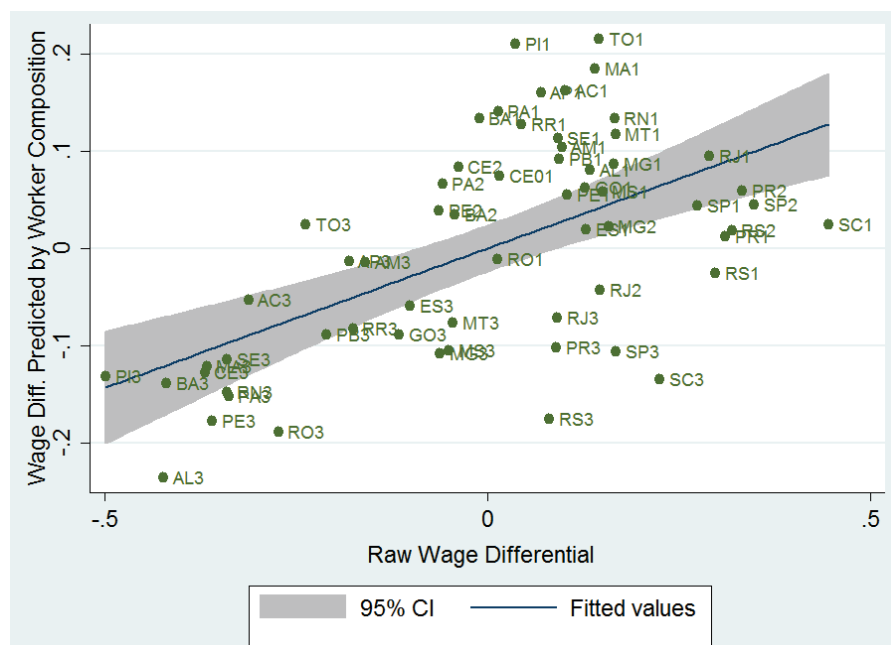


FIGURE 1
Wage levels across localities - Composition effect



FIGURE 2
Wage levels across localities - Minority

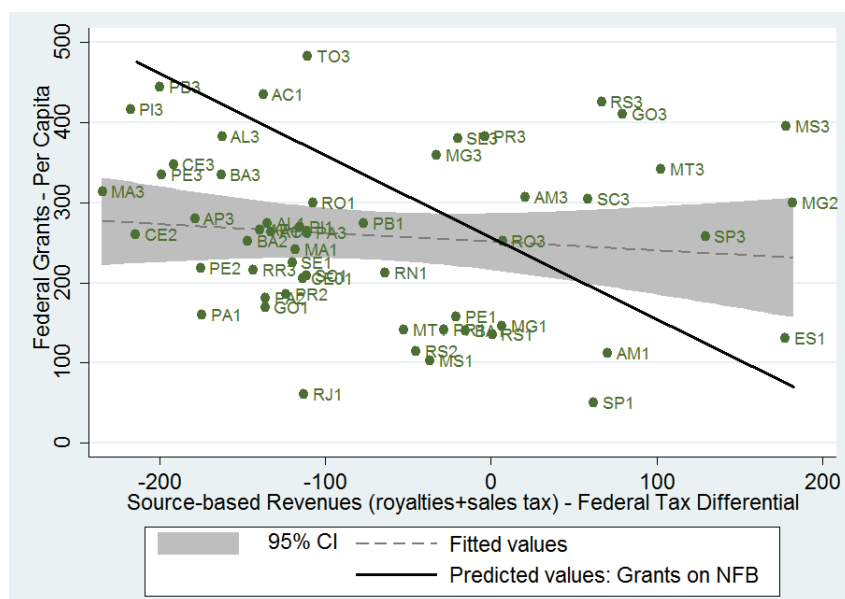


FIGURE 3
Grants relative to Federal taxes and source-based revenues (per capita)

Supporting Information

Appendix A.1

We estimate composition and location effects using our Function (1) with individual microdata extracted from the PNAD dataset. The household sample is different for each year; therefore, we run separate regressions yearly in accordance with Function (1). We use information on individual characteristics, such as education, gender, age, ethnic group, immigrant status (lives in the city in which he/she was born), and occupation, such as weekly working hours, whether the individual works in a formal market and whether the individual is self-employed. We interact most indicators with gender. The estimated composition wage is obtained multiplying each coefficient (estimated on yearly individual data) by the location average of the respective independent variable. As in Albouy (2012), the average wage level in each locality follows the sum of the location and composition wage. In accordance with Function (1), the location wage is obtained by the difference between the fitted wage and the predicted worker composition. The descriptive analysis (average from 2004 to 2009) on independent variables by locality is available in Table A.1 below. We split independent variables based on the use to estimate location and composition wage, respectively. Coefficient estimates on wage for our five year period of analysis are available in Tables A.2.

Appendix A.2

In Albouy's model, the sum of these three components plus intergovernmental grants provide the communities' Net Fiscal Benefit (NFB). Thus, Function (1) provides the relation between grants (F) and the three terms of fiscal components to evaluate a fiscal policy in terms of efficiency and equalization:²¹

$$\bar{F}^j = \tau_W^F(\bar{w}^j - \bar{w}^{j,F}) - \frac{\tau_L^j r^j L^j + \tau_K^j i^j K^j}{N^j} + (1 - \alpha) \frac{p_G^i G^j}{N^j} + \bar{F}_e^j \quad (1)$$

Where:

Households provide labor markets with different types of individuals (different skills) indexed by $e = 1, \dots, E$; who can live in any locality J , indexed by $j = 1, \dots, J$. Population composition on locality j follows: $N^j = (N_1^j, \dots, N_E^j)$. Sourced based taxes on land and capital, plus destination based taxes on wages are represented by $\tau_L^j, \tau_K^j, \tau_W^j$; respectively. Local communities expenses for public goods follow the term $p_G^i G^j$. Moreover, the term $(1-\alpha)$ designates public goods externalities (e. g. in the case in which individuals contributes to the public goods in a ratio larger than they consume them, we have $\alpha < 1$). The mean of grants that region F obtains from central government follows:

$$\bar{F}^j = \frac{1}{N^j} \sum_e N_e^j F_e^F$$

As emphasized by Albouy (2012), a larger positive federal-tax differential drives wage employees to migrate to regions where the fiscal burden is smaller and the net fiscal benefit is larger, considering that individuals pay for the former and benefit from the latter. Based on this framework, efficiency policies should refund workers that hold a tax burden larger than average to avoid tax differential distortions in geographic location choices. In short, inefficient redistribution of resources distorts migration decisions.²² The same idea can be used to address source-based tax, because the tax price should be associated with the public provision local cost.

Re-arranging terms in function (1), one obtains:

$$\bar{F}^j - \tau_W^F(\bar{w}^j - \bar{w}^{j,F}) + \frac{\tau_L^j r^j L^j + \tau_K^j i^j K^j}{N^j} = (1 - \alpha) \frac{p_G^i G^j}{N^j} + \bar{F}_e^j \quad (2)$$

Where the terms in the left summarize the effect of the NFB. It is important to note that all terms on the left can be estimated empirically. This relation, as developed by Albouy (2012), is the basis for our methodological approach detailed in Section 4.²³ In this approach, an efficient grant policy should work to mitigate larger fiscal needs differences. In fact, the optimal level of grants (term F_e^{Fj} in function 2) considers the same level of NFB between regions.

²¹ For the model's complete set-up and description, refer to Albouy (2012) pp. 825-827.

²² Most models assume that governments ignore the effect of their decisions over migration.

²³ In Boadway and Flatters (1982) for example, the net fiscal benefit follows source based tax differential and net fiscal benefit supported by federal income tax.

Table A.1

Coefficients estimates on Wage - Cross Section

Dependent variable: Ln wage		2005	2006	2007	2008	2009
Industry sector	Gender (D=1 if female)	0.090 (0.486)	-0.226 (-1.129)	0.140 (0.612)	-0.160 (-0.678)	0.142 (0.603)
	Age (Ln)	0.811*** (26.709)	0.778*** (23.362)	0.826*** (22.305)	0.791*** (20.519)	0.693*** (17.697)
	Ethnic group (D=1 if white)	0.251*** (15.050)	0.277*** (15.646)	0.271*** (14.142)	0.292*** (14.717)	0.270*** (13.016)
	Migrant (D=1 if lives in na city different from where was bom)	0.045*** (2.691)	0.037** (2.092)	-0.048** (-2.475)	0.009 (0.457)	0.062*** (2.998)
	Schooling (Ln)	0.457*** (38.075)	0.455*** (37.454)	0.470*** (35.666)	0.439*** (30.636)	0.447*** (30.827)
	Agriculture (D=1, zero otherwise)	-0.424*** (-11.898)	-0.167*** (-5.734)	-0.083** (-2.530)	-0.388*** (-9.200)	-0.479*** (-11.338)
	Retail Trade (D=1, zero otherwise)	-0.275*** (-8.796)	0.076*** (2.764)	0.084*** (2.933)	-0.151*** (-4.125)	-0.222*** (-5.901)
	Construction (D=1, zero otherwise)	-0.287*** (-8.110)	.	.	-0.219*** (-5.387)	-0.283*** (-6.788)
	Manufacturing (D=1, zero otherwise)	-0.314*** (-10.110)	0.041 (1.501)	0.039 (1.346)	-0.197*** (-5.451)	-0.240*** (-6.361)
	Government (D=1, zero otherwise)	.	0.336*** (8.908)	0.403*** (9.999)	.	.
	Education and Health Care (D=1, zero otherwise)	-0.295*** (-9.242)	0.024 (0.712)	0.047 (1.304)	-0.285*** (-7.871)	-0.309*** (-8.032)
	Others	-0.214*** (-7.274)	0.172*** (6.396)	0.191*** (6.756)	-0.089*** (-2.662)	-0.143*** (-4.193)
	Informal sector job	-0.537*** (-15.955)	-0.587*** (-15.545)	-0.530*** (-13.227)	-0.544*** (-13.125)	-0.495*** (-11.841)
	Formal sector job	-0.225*** (-7.009)	-0.287*** (-7.781)	-0.245*** (-6.533)	-0.345*** (-9.288)	-0.253*** (-6.469)
	Self-employed	-0.339*** (-9.665)	-0.370*** (-9.242)	-0.217*** (-5.232)	-0.222*** (-5.053)	-0.218*** (-4.980)
	Working time less than 40 hours weekly (D=1, zero otherwise)	.	.	-0.375*** (-17.771)	-0.366*** (-16.552)	-0.423*** (-18.977)
	Working time between 40-48 hours weekly (D=1, zero otherwise)	0.381*** (21.562)	0.415*** (22.518)	-0.008 (-0.506)	0.027 (1.631)	0.033* (1.911)
	Working time more than 48 hours weekly (D=1, zero otherwise)	0.406*** (22.758)	0.438*** (24.222)	.	.	.
	Mother with any child under 14 age (D=1, zero otherwise)	-0.079*** (-3.424)	-0.014 (-0.535)	-0.025 (-0.798)	0.001 (0.040)	-0.058* (-1.791)
	Executive or managerial occupations (D=1, zero otherwise)	0.824*** (23.945)	0.951*** (24.624)	0.986*** (23.912)	1.029*** (24.579)	1.039*** (23.885)
Interactions	Ethnic group (D=1 if white) * Gender (D=1 if female)	0.008 (0.315)	-0.015 (-0.587)	-0.035 (-1.214)	-0.011 (-0.357)	0.023 (0.740)
	Age (Ln) * Gender (D=1 if female)	-0.180*** (-3.839)	-0.123** (-2.439)	-0.193*** (-3.334)	-0.120** (-2.053)	-0.213*** (-3.601)
	Migrant * Gender (D=1 if female)	-0.046* (-1.846)	-0.034 (-1.305)	-0.001 (-0.021)	-0.022 (-0.746)	-0.055* (-1.778)
	Schooling (Ln) * Gender (D=1 if female)	0.005 (0.298)	0.051*** (2.791)	0.013 (0.616)	0.015 (0.648)	0.024 (1.096)
	Informal sector job * Gender (D=1 if female)	0.060 (1.421)	0.065 (1.408)	0.032 (0.639)	0.066 (1.272)	0.012 (0.227)
	Formal sector job * Gender (D=1 if female)	0.237*** (5.638)	0.285*** (6.207)	0.266*** (5.501)	0.259*** (5.486)	0.251*** (5.035)
	Self-employed * Gender (D=1 if female)	-0.027 (-0.551)	-0.029 (-0.557)	-0.074 (-1.320)	-0.036 (-0.594)	0.028 (0.464)
	Executive or managerial occupations * Gender (D=1 if female)	0.106** (2.135)	-0.011 (-0.214)	-0.102* (-1.808)	-0.048 (-0.828)	0.052 (0.867)
	Adjusted R ²	0.101	0.096	0.076	0.068	0.068
	Number of observations	143,681	147,165	144,738	146,385	150,116

Obs: In parentheses are robust t-statistics. *Significant at 10%; **significant at 5%; ***significant at 1%.

Table A.2

Descriptive Statistics (Mean) - Cross Section	2005	2006	2007	2008	2009
Dependent variable: Ln wage	6.60	6.70	6.81	6.89	6.95
Gender (D=1 if female)	0.41	0.41	0.41	0.42	0.42
Age (Ln)	3.57	3.57	3.57	3.58	3.58
Ethnic group (D=1 if white)	0.48	0.48	0.47	0.46	0.46
Migrant (D=1 if lives in na city different from where was born)	0.53	0.52	0.51	0.51	0.50
Schooling (Ln)	2.00	1.88	1.90	1.93	1.95
Agriculture (D=1, zero otherwise)	0.11	0.10	0.10	0.10	0.10
Retail Trade (D=1, zero otherwise)	0.19	0.18	0.19	0.18	0.18
Construction (D=1, zero otherwise)	0.07	0.08	0.08	0.08	0.08
Manufacturing (D=1, zero otherwise)	0.15	0.15	0.15	0.15	0.15
Government (D=1, zero otherwise)	0.06	0.06	0.06	0.06	0.06
Education and Health Care (D=1, zero otherwise)	0.11	0.11	0.11	0.11	0.11
Others	0.30	0.31	0.31	0.32	0.32
Informal sector job	0.24	0.23	0.23	0.22	0.22
Formal sector job	0.36	0.37	0.38	0.39	0.40
Self-employed	0.24	0.24	0.24	0.22	0.22
Working time less than 40 hours weekly (D=1, zero otherwise)	0.21	0.22	0.21	0.21	0.21
Working time between 40-48 hours weekly (D=1, zero otherwise)	0.37	0.38	0.41	0.42	0.44
Working time more than 48 hours weekly (D=1, zero otherwise)	0.41	0.40	0.38	0.37	0.35
Mother with any child under 14 age (D=1, zero otherwise)	0.06	0.06	0.05	0.05	0.05
Executive or managerial occupations (D=1, zero otherwise)	0.13	0.14	0.14	0.14	0.14
Ethnic group (D=1 if white) * Gender (D=1 if female)	0.21	0.21	0.21	0.20	0.20
Age (Ln) * Gender (D=1 if female)	1.45	1.47	1.48	1.49	1.51
Migrant * Gender (D=1 if female)	0.22	0.22	0.21	0.22	0.21
Schooling (Ln) * Gender (D=1 if female)	0.88	0.83	0.84	0.86	0.88
Informal sector job * Gender (D=1 if female)	0.12	0.12	0.11	0.11	0.11
Formal sector job * Gender (D=1 if female)	0.13	0.13	0.14	0.14	0.14
Self-employed * Gender (D=1 if female)	0.08	0.08	0.08	0.08	0.08
Executive or managerial occupations * Gender (D=1 if female)	0.07	0.07	0.07	0.07	0.07

Transparência fiscal, comprometimento do governo e credibilidade fiscal:

Uma análise considerando os efeitos da contabilidade criativa no Brasil

Tatiana Acar

Departamento de Economia, Universidade Federal Fluminense (UFF), Rio de Janeiro - Brasil.
tatiacar@hotmail.com

Resumo

Desde 2009, o governo brasileiro faz uso de contabilidade criativa com intuito de melhorar os resultados relativos ao superávit primário. Com isso, a transparência fiscal vem se reduzindo e a credibilidade fiscal deteriorando. Este estudo tem como principal objetivo analisar o efeito da transparência das contas públicas sobre a credibilidade fiscal no Brasil. Dentre as novidades do estudo estão: (i) a apresentação de um novo índice de credibilidade fiscal, e (ii) a apresentação de um indicador de transparência fiscal, obtido por meio de um valor mais realista para o superávit primário, o qual capta o real esforço de economia do governo desconsiderando os efeitos da contabilidade criativa do governo. Os achados sugerem que melhorias artificiais do resultado fiscal por meio de contabilidade criativa reduzem a credibilidade fiscal. Os resultados também reforçam a importância da solvência fiscal e do comprometimento com metas fiscais para a construção de credibilidade. Ademais, estimações por regressão quantílica apontam importantes achados no tocante à mudança na magnitude dos efeitos das principais variáveis analisadas (como a transparência e o comprometimento fiscal), quando considerados diferentes níveis de credibilidade fiscal.

Palavras-chave: transparência fiscal; credibilidade fiscal; comprometimento; contabilidade criativa

Abstract

Since 2009, the Brazilian government makes use of creative accounting in order to improve the results for the primary surplus. Thus, fiscal transparency has been reduced and fiscal credibility deteriorating. This study aims to analyze the effect of the transparency of public accounts on fiscal credibility in Brazil. Among the novelties of the study are: (i) the presentation of a new fiscal credibility index, and (ii) the presentation of an indicator of fiscal transparency achieved through a more realistic value for the primary surplus, which captures the real effort of government savings excluding the effects of creative accounting. The findings suggest that artificial improvements of the fiscal balance through creative accounting reduce fiscal credibility. The results also reinforce the importance of fiscal solvency and commitment to fiscal targets for building credibility. Moreover, quantile regression estimates show important findings with respect to changes in the magnitudes of the effects of key variables (such as transparency and fiscal commitment) when considered different levels of fiscal credibility.

Keywords: fiscal transparency; fiscal credibility; commitment; creative accounting

Área 4: Macroeconomia Aplicada (Applied Macroeconomics)

Classificação JEL: E62, H62, H63

1. Introdução

Desde o final de 2009, quando países da União Européia passaram a apresentar sérios problemas de solvência fiscal, a credibilidade fiscal tem sido amplamente citada como um dos fundamentos mais importantes da política macroeconômica. Neste contexto, a transparência fiscal ganha importância, pois torna mais eficaz a formulação de políticas e o gerenciamento dos riscos fiscais, além de garantir que os governos forneçam ao público informações mais acuradas (FMI, 2012)¹. Não obstante, a opacidade fiscal decorrente de contabilidade criativa está se popularizando e virando prática disseminada ao redor do mundo (Schettini et al., 2011; Petersen, 2003; Journard, 2008). A motivação para tais práticas varia, sendo desde o incremento dos resultados com fins eleitorais, bem como para atingir determinada meta devido a regras fiscais rígidas, ou a redução do controle parlamentar ao ocultar gastos que precisariam de prévia aprovação do congresso.²

Assegurar a solvência fiscal é importante para a conquista de credibilidade fiscal e atração de recursos financeiros e investimentos privados (Blanchard, 2010). Evidências empíricas apontam que o comprometimento do governo com as contas públicas e a credibilidade fiscal afetam as expectativas e a confiança dos empresários e consumidores, como também o desempenho macroeconômico (de Mendonça e Machado, 2013; de Mendonça e Tostes, 2015; Montes e Acar, 2015). Um país com credibilidade fiscal elevada possui maior capacidade de atração de investimentos privados, consegue alavancar mais recursos para financiar os investimentos públicos e a dívida pública, e consegue aumentar a eficácia das políticas macroeconômicas e, assim, propiciar um ambiente favorável para o crescimento econômico. Ademais, a credibilidade fiscal mais elevada reduz a possibilidade de mudanças abruptas nas expectativas e confiança dos investidores e, por conseguinte, reduz a vulnerabilidade externa.

Acadêmicos e policymakers reconhecem que a transparência na gestão das contas públicas é uma précondição para a redução da volatilidade dos mercados e para a sustentabilidade fiscal. Evidências empíricas apontam para uma relação positiva entre o grau de transparência e a percepção do mercado sobre a solvência fiscal. Peat et al (2015) e Glennerster e Shin (2008) encontram que uma maior opacidade fiscal leva a maiores spreads no mercado de crédito soberano. Já Arbatli e Escolano (2012) associam a transparência fiscal a melhores notas de risco soberano, e o estudo de Bernoth e Wolff (2008) constata que a contabilidade criativa na União Européia – a qual reduz a transparência – elevou o prêmio de risco na região da zona do euro. Sem transparência, as variáveis fiscais oficiais divulgadas pelo governo podem não fornecer ao público resultados confiáveis e que garantam a correta previsibilidade em relação à verdadeira situação fiscal do país, o que causa danos à credibilidade e sua efetividade. (Kopits e Craig, 1998).

A consequente falta de transparência está relacionada à tentação dos governantes em manipular os resultados fiscais por meio de contabilidade criativa. Embora incipiente, a literatura que analisa as interferências do governo nas contas públicas por meio de contabilidade criativa tem avançado: existem estudos para a União Européia (Koen e Noord, 2005; Von Hagen e Wolff, 2006), para países da OCDE (Journard et. al,

¹ Segundo o FMI (2012), a última década e meia representou esforços de promoção de normas internacionais para garantir a transparência.

² Alt e Lassen (2006) e Von Hagen e Wolff (2006) mostram que a contabilidade criativa varia de acordo com os ciclos eleitorais. Estudos sobre as implicações de regras fiscais rígidas nos resultados fiscais podem ser vistos em Bunch (1991), von Hagen (1991), Kiewiet e Szakaly (1996).

2008) e para os Estados Unidos (Petersen, 2003). Por outro lado, a literatura indica que regras fiscais rígidas aumentam a tentação do governo em se beneficiar do grau de liberdade gerado pelas convenções contábeis. Isto pode ocorrer quando o cumprimento das metas fiscais se torna mais difícil, com o consequente aumento da volatilidade macroeconômica e perda de credibilidade fiscal (Milesi-Ferretti, 2004; Ter-Minassian, 2010).

Por sua vez, o desafio de determinar se um governo possui margem para levantar recursos adicionais para gastos meritórios tem confrontado o FMI nos últimos anos. Diante da inevitabilidade dos ajustes fiscais devido à crise da dívida soberana dos países europeus em 2009, questões relacionadas ao nível ótimo de gastos adicionais (como investimento em infra-estrutura e gastos em programas sociais) que não comprometa a sustentabilidade fiscal se tornaram frequentes. Assim, atenções se direcionaram para o grau em que um país possui “espaço-fiscal” (*fiscal space*³), ou em outras palavras, se um país é capaz de realizar dispêndios em setores desejados sem comprometer a sustentabilidade da dívida. Em estudo sobre consolidação fiscal e investimento em países latino-americanos, Carranza et. al (2014) conclui que altos investimentos em infra-estrutura devem vir acompanhados de melhores processos de gasto, viabilizados por um “espaço fiscal” suficientemente alto. Estudos empíricos utilizam as variáveis de *fiscal space* propostas por Aizenman e Jinjark (2010) para avaliar seus efeitos no risco soberano, utilizando como *proxy* para o risco o *credit default swap* (CDS) (Aizenman et. al, 2013; Peat et. al, 2015).

Apesar dos avanços na literatura, existem lacunas a se preencherem quanto à mensuração da credibilidade e da transparência fiscal e à necessidade de evidências empíricas acerca dos determinantes da credibilidade fiscal. Em particular, existe a necessidade de evidências empíricas acerca dos efeitos da transparência fiscal e do comprometimento do governo com as contas públicas sobre a credibilidade fiscal. Nesse sentido, este estudo busca preencher essas lacunas e tem como principais objetivos: (i) desenvolver um novo índice de credibilidade fiscal, (ii) captar os efeitos da contabilidade criativa nas contas do governo e, assim, observar o nível de transparência fiscal, e (iii) identificar os determinantes da credibilidade fiscal, com ênfase para os efeitos da transparência fiscal e do *fiscal space*.

A despeito do esforço empreendido por de Mendonça e Machado (2013) na construção de um indicador de credibilidade fiscal, o indicador proposto por eles não se mostra suficiente para explicar o atual cenário de deterioração fiscal do Brasil, visto que a variável de expectativa para a dívida pública não capta a situação fiscal de forma integral. Por outro lado, são poucos os estudos que abordam os efeitos da transparência fiscal na economia, e devido à dificuldade de mensuração, são poucos os estudos que analisam os efeitos na economia oriundos da contabilidade criativa nas contas do governo.

Este artigo analisa os efeitos que exercem a transparência na gestão das contas públicas e o compromisso com a estabilidade fiscal sobre a credibilidade fiscal. O estudo sugere que, se por um lado, interferências nos resultados fiscais por meio de contabilidade criativa geram resultados maiores para o superávit primário mesmo com expansão dos gastos (possibilitando, assim, expandir ainda mais os gastos e a dívida pública), por outro lado, essas manobras contábeis minam a transparência fiscal e, por conseguinte a credibilidade fiscal. Ou seja, quando a transparência fiscal é reduzida, os resultados fiscais divulgados pelo governo deixam de ser confiáveis e não garantem a

³ O termo *fiscal space* é relativamente novo na literatura, e as variáveis assim denominadas buscam captar diferentes medidas de sustentabilidade das finanças públicas. A medida de *fiscal space* construída e utilizada na análise empírica está baseada em Aizenman e Jinjark (2010).

correta previsibilidade em relação à verdadeira situação fiscal do país, causando, nesse sentido, danos à credibilidade fiscal. Destarte, uma das questões centrais colocada no presente estudo diz respeito ao efeito dessas práticas sobre a transparência fiscal e, por conseguinte, sobre a credibilidade fiscal.

Como principais resultados, o trabalho encontra evidências empíricas de que a credibilidade fiscal se deteriorou nos últimos anos, sugerindo que a falta de transparência é percebida – e mal vista – pelo público. Ou seja, os resultados fiscais criados por meio de contabilidade criativa reduzem a confiança que o público deposita no governo e, assim, deterioram a credibilidade fiscal, produzindo um efeito contrário ao pretendido pelo governo. Os resultados reforçam a relevância da solvência fiscal para a construção da credibilidade fiscal. Além disso, estimações por regressão quantílica apontam importantes achados no tocante à mudança na magnitude dos efeitos das principais variáveis analisadas, quando considerados diferentes níveis de credibilidade fiscal.

Portanto, o presente estudo se diferencia e contribui para a literatura sobre credibilidade e transparência fiscal nos seguintes aspectos. Em primeiro lugar, desenvolvemos um novo índice de credibilidade fiscal, o qual aperfeiçoa o índice proposto por de Mendonça e Machado (2013). Em segundo, elaboramos um indicador de transparência fiscal, o qual captura o grau de interferência do governo nos resultados relacionados ao superávit primário por meio de contabilidade criativa. Em terceiro, analisamos os determinantes da credibilidade fiscal, com ênfase para os efeitos da transparência fiscal, do comprometimento do governo com os seus objetivos e da solvência (por meio da variável “*fiscal space*”). A análise é desenvolvida para o Brasil e compreende o período entre Dezembro de 2002 e Novembro de 2014. O Brasil é um interessante estudo de caso, pois com a adoção do regime de metas para inflação, o governo se viu obrigado a implantar políticas de ajuste das contas públicas e de gerenciamento da dívida de modo a desenvolver sua credibilidade. Entretanto, desde o final do ano de 2009, o governo tem se afastado das metas fiscais e feito uso de artifícios contábeis em suas contas. Essas ações estão gerando consequências adversas em termos de formação de expectativas dos agentes, desempenho macroeconômico e instabilidade política. O cenário adverso que se materializou, tanto na economia quanto na política, coloca em risco o regime de metas para inflação e a própria permanência da Presidente da República Dilma Rousseff em seu cargo, a qual sofre processo de *impeachment* por crimes de responsabilidade.

2. Evidências empíricas para a economia brasileira

Desde o final da década de 1990, percebe-se no Brasil uma mudança gradual na condução da política fiscal, e esforços por parte do governo em melhorar a disciplina fiscal.⁴ Esses esforços foram ampliados com a adoção do regime de metas para inflação (RMI) em junho de 1999, pois um dos pilares desse regime é a não existência de dominância fiscal. No caso brasileiro, a melhora do compromisso do governo com a disciplina fiscal foi reforçada pela Lei de Responsabilidade Fiscal (LRF) criada em 2000, e com a adoção das metas de superávit primário. A ideia é que com a adoção

⁴ No final de 1998, o país precisou recorrer a US\$ 41 bilhões em empréstimos do Fundo Monetário Internacional (FMI), Banco Mundial, Banco Interamericano de Desenvolvimento (BID) e o Banco de Compensações Internacionais (BIS) devido a desequilíbrios fiscais, e como um dos requisitos, foi obrigado a assegurar uma maior disciplina fiscal.

dessas medidas fosse observada a melhora do desempenho macroeconômico e, assim, da credibilidade do regime de metas para inflação.⁵

A despeito de mudanças estruturais no que tange à disciplina fiscal, o cenário fiscal tornou-se incerto a partir de 2009 com a queda abrupta da atividade econômica devido ao agravamento da crise internacional. Nesse contexto, o governo passa a adotar de forma recorrente medidas para obter receitas fiscais atípicas, além de postergar pagamentos e aumentar de forma significativa o passivo com bancos públicos referentes aos atrasos dos repasses de recursos próprios que os bancos públicos utilizam para financiar os programas sociais (as chamadas “pedaladas fiscais”).⁶ Neste mesmo ano houve uma reversão das expectativas formadas em relação ao superávit primário para valores abaixo da meta observada, e grande parte dessa deterioração de expectativas pode ser explicada pelas ações tomadas pelo governo no tocante à gestão das contas públicas. O governo passa, também, a reduzir a meta de superávit primário em vigor por meio de alterações na lei orçamentária do ano (em alguns casos a alteração era feita no final do exercício financeiro, inviabilizando as metas de superávit) (ver tabela A.1, no Apêndice). Além disso, a partir de meados de 2013, a taxa básica de juros Selic inicia novo ciclo de elevação após ter atingido seu menor valor da série histórica e passa a pressionar a dívida pública.

As ações tomadas pelo governo e o cenário econômico em deterioração levaram a superávits primários cada vez menores a cada período. Devido à deterioração do cenário macroeconômico e à incapacidade do governo em reverter a situação fiscal, a confiança dos investidores em relação ao Brasil diminuiu (o Brasil passa a ter rebaixamentos de notas pelas principais agências de rating desde o final de 2015). Além disso, a credibilidade fiscal inicia trajetória de deterioração (a Figura 1 mostra esse comportamento para a credibilidade fiscal, o qual é captado pelo indicador proposto no presente estudo).

Apesar do progresso na literatura referente ao lado fiscal, há lacunas a serem preenchidas quanto à mensuração da credibilidade fiscal, e necessidade de estudos empíricos sobre seus determinantes.

2.1 Dados

O índice de credibilidade fiscal proposto por de Mendonça e Machado (2013) representa um esforço no que diz respeito ao desenvolvimento de indicadores capazes de medir a credibilidade fiscal. Contudo, o índice pode não refletir de forma fiel a avaliação dos agentes sobre a situação fiscal como um todo, pois o índice apenas capta as expectativas formadas em relação à dívida líquida. Essas expectativas podem estar baixas, mas não necessariamente o governo está sendo bem avaliado pelos investidores com relação aos seus compromissos fiscais de modo geral. Por exemplo, sua capacidade de arrecadação de impostos e de honrar os compromissos pode estar comprometida, o que gera suspeita sobre a estabilidade fiscal do país. Sob esta ótica, o referido índice não considera outros aspectos que podem influenciar a percepção dos investidores sobre o

⁵ Em suma, a LRF define limites de endividamento e controle de gastos excessivos para as três esferas de governo, e previne que o Governo Federal financie os estados e municípios.

⁶ O objetivo dessas medidas é o de inflar artificialmente o resultado primário e facilitar a tarefa do gestor público, pois serão menores os seus esforços para cumprir rigorosamente a legislação, as obrigações fiscais, assim como as metas de superávit primário estabelecidas. Com isso o governo consegue – no curto prazo – assegurar uma boa imagem perante seu eleitorado, e driblar a desconfiança e fiscalização rigorosa dos órgãos fiscalizadores.

compromisso fiscal do governo. Em outras palavras, o indicador é menos sensível ao ambiente fiscal como um todo.

Buscando atenuar esta vulnerabilidade do indicador proposto por de Mendonça e Machado (2013) e, assim, contribuir com a literatura, o presente estudo propõe um aperfeiçoamento a partir da inclusão de dois componentes ao índice: o desvio da expectativa de superávit primário em relação à meta, e um redutor que capta as situações em que a meta de superávit primário está abaixo do superávit primário necessário para manter a dívida sustentável. Assim, quando o governo não assume uma postura de compromisso com as contas públicas, é penalizado por não fazer convergir as expectativas formadas para o superávit primário em direção à meta anunciada. Além disso, a decisão do governo por metas de superávit abaixo do necessário para manter estável a dívida pública também afeta a confiança dos agentes em relação à sustentabilidade da dívida e reduz a credibilidade fiscal. Nesse sentido, o índice de credibilidade fiscal (fc) é construído da seguinte maneira:

$$(1) \quad fc = \begin{cases} 1 & \text{if } E_t(debt_{t+12}) \leq debt_t^{Min} \\ 1 - \left\{ \frac{[E_t(debt_{t+12}) - debt_t^{Min}]}{debt_t^{Max} - debt_t^{Min}} \right\} & \text{if } debt_t^{Min} < E_t(debt_{t+12}) < debt_t^{Max} \\ 0 & \text{if } E_t(debt_{t+12}) \geq debt_t^{Max} \end{cases} X \left\{ \begin{array}{ll} \frac{E_t(PS_{t+12})}{PS_target_{t+12}}, & \text{if } \frac{E_t(PS_{t+12})}{PS_target_{t+12}} < 1 \\ 1, & \text{if } \frac{E_t(PS_{t+12})}{PS_target_{t+12}} \geq 1 \end{array} \right\} X \left\{ \begin{array}{ll} \frac{PS_target_{t+12}}{PS_req_{t+12}}, & \text{if } \frac{PS_target_{t+12}}{PS_req_{t+12}} < 1 \\ 1, & \text{if } \frac{PS_target_{t+12}}{PS_req_{t+12}} \geq 1 \end{array} \right\}$$

Em que, $E_t(debt_{t+12})$ representa a expectativa formada no instante t para a dívida 12 períodos à frente, $debt_t^{Max}$ e $debt_t^{Min}$ representam os limites superior e inferior da dívida, $E_t(PS_{t+12})$ representa a expectativa formada no instante t para o superávit primário 12 períodos à frente, PS_target_{t+12} representa a meta para o superávit primário 12 períodos à frente e PS_req_{t+12} representa o superávit primário requerido para manter estável a dívida pública 12 períodos à frente. Por sua vez, o superávit primário requerido é calculado da seguinte maneira: $PS_req = Debt_{t-1} \cdot \left[\frac{(i-g)}{(1-g)} \right]$; tal que g representa a taxa de crescimento do produto e i é a taxa de juros.⁷

O primeiro termo representa o índice de credibilidade fiscal tal como proposto por de Mendonça e Machado (2013), em que os limites inferior ($debt_t^{Min}$) e superior ($debt_t^{Max}$) referem-se a uma dívida de 40% do PIB⁸ e 60% do PIB⁹, respectivamente. O segundo termo $\left[\frac{E_t(PS_{t+12})}{PS_target_{t+12}} \right]$ representa a razão entre o superávit primário esperado e a meta de superávit primário anunciada, e mostra o grau de confiança do público no anúncio realizado. Quando $\left[\frac{E_t(PS_{t+12})}{PS_target_{t+12}} \right] < 1$, significa que os analistas não acreditam que o governo irá cumprir a meta e, assim, funciona como um redutor do índice de credibilidade fiscal. Por outro lado, quando $\left[\frac{E_t(PS_{t+12})}{PS_target_{t+12}} \right] \geq 1$, o governo sinaliza que é comprometido com a meta de superávit primário, e não há impacto negativo na

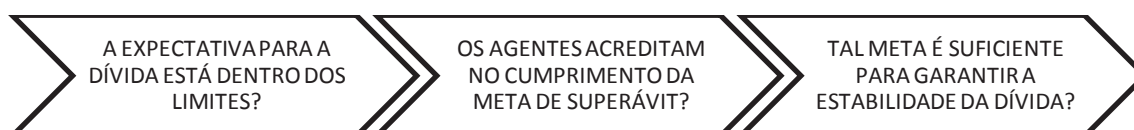
⁷ A medida de PS_req é amplamente conhecida. Ver, por exemplo, Celasun, Debrun e Ostry (2006) e Levy (2010).

⁸ Tal limite está de acordo com a recomendação do FMI (2002) sobre o objetivo de longo prazo que não deve ser quebrado pelas economias emergentes.

⁹ Uma razão dívida/PIB de 60% é a meta estabelecida pelo Tratado de Maastricht, sendo, portanto, uma referência aos países em desenvolvimento.

credibilidade fiscal. O terceiro termo demonstra se o esforço fiscal a ser realizado pelo governo (dado pela meta de superávit) irá assegurar uma dívida estável. Deste modo, o termo $\left[\frac{PS_target_{t+12}}{PS_req_{t+12}} \right]$ significa o desvio da meta de superávit primário em relação ao que seria necessário para manter a dívida estável ao longo do tempo. Quando a meta de superávit primário definida pelo governo é menor do que a meta requerida, $\left[\frac{PS_target_{t+12}}{PS_req_{t+12}} \right] < 1$, é sinalizado ao público um baixo esforço de poupança comparado ao necessário para manter a dívida inalterada, e assim, a credibilidade fiscal é penalizada. Por outro lado, quando $\left[\frac{PS_target_{t+12}}{PS_req_{t+12}} \right] \geq 1$, não há punição.

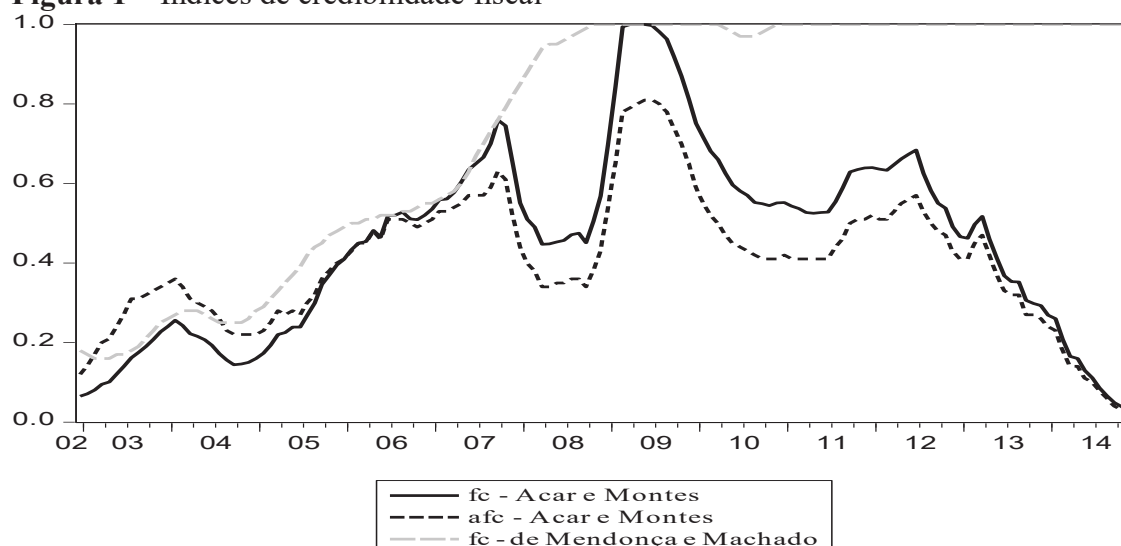
Deste modo, as três questões que se colocam para avaliar se o país dispõe de credibilidade – a partir do índice proposto – são:



A fim de conferir robustez ao índice e evitar dúvidas referentes aos limites utilizados no indicador, uma análise de sensibilidade em relação à sua forma funcional foi construída ao alterar os parâmetros subjacentes, dando origem ao Índice de credibilidade ajustado (*afc*). O limite superior da dívida é ampliado para 70% do PIB, enquanto o inferior é reduzido para 30% do PIB. Assim, é mais difícil obter credibilidade máxima, ao passo que é também mais difícil perder completamente a credibilidade.

A Figura 1 mostra os gráficos para as trajetórias dos índices de credibilidade fiscal ao longo do tempo.

Figura 1 – Índices de credibilidade fiscal



Fonte: Gráfico elaborado pelos autores.

Comparando os índices de Acar e Montes com o índice elaborado por de Mendonça e Machado (2013), é possível observar que a partir de 2008 há uma dissociação no comportamento dos indicadores, pois se inicia uma deterioração das expectativas de superávit primário em relação à meta. Além disso, em 2013, grandes reduções na meta de superávit primário também contribuíram para deteriorar a

credibilidade fiscal captada por *fc* e *afc* (Acar e Montes). Se por um lado, metas baixas facilitam a tarefa do governo em convencer os agentes de que a meta será alcançada, por outro lado, metas muito baixas indicam uma situação de fragilidade fiscal, em que o governo enfrenta dificuldades de arrecadação de impostos e/ou apresenta gastos excessivos, e levam a deterioração da confiança dos agentes.

Os dados utilizados no estudo possuem frequência mensal, contemplando o período entre dezembro de 2002 e novembro de 2014.¹⁰

No tocante à contabilidade criativa, Koen e Noord (2005) identificam três principais conceitos relacionados às operações atípicas que inflam as estatísticas fiscais e não possuem relação com o fortalecimento permanente da posição fiscal do governo: (i) trapaças fiscais (*fiscal gimmicks*) – essas apenas modificam as estatísticas fiscais de maneira “fictícia” e/ou são compensadas por um efeito contrário no exercício fiscal seguinte, como por exemplo, a superestimação e antecipação de receitas ou postergação de pagamentos; (ii) contabilidade criativa – na qual se utiliza das flexibilidades e omissões existentes nas normas contábeis para conferir tratamento pouco usual a operações intra-setor público¹¹, e; (iii) medidas extraordinárias (*one-off measures*) – decisões governamentais de natureza não recorrente, que afetam a posição fiscal contemporânea e não se mantêm no futuro. Estas medidas podem se tornar nocivas quando se contrai despesas obrigatórias baseando-se em receitas que não se constituem fontes garantidas no futuro.

A fim de observar o impacto da falta de transparência fiscal sobre a credibilidade fiscal, uma nova variável é criada para medir o grau de interferência do governo nos resultados relacionados ao superávit primário por meio de contabilidade criativa. A variável chamada *transparency* é construída pela razão entre o superávit realista (*real_PS*) e o superávit divulgado pelo governo (*released_PS*), ou seja, $transparency = \frac{real_PS}{released_PS}$. O superávit realista corresponde ao superávit oficial divulgado deduzido das interferências contábeis artificiais (contabilidade criativa), e busca capturar o verdadeiro esforço de poupança do governo.

São consideradas como contabilidade criativa aquelas operações que não afetam o superávit primário, tampouco a dívida líquida, porém aumentam o endividamento, e assim, não possuem relação com o fortalecimento permanente da posição fiscal do governo, constituindo-se em mecanismos contábeis pouco transparentes. Seguindo linhas gerais do tratamento dado na literatura para outros países¹², tomou-se como critério para a classificação dos eventos em contabilidade criativa o atendimento de um ou mais dos seguintes quesitos: i) efeito sobre o patrimônio líquido do setor público, ii) perspectivas de repetição no longo prazo (recorrência); iii) indisponibilidade de informações públicas sobre a operação. Assim, as seguintes naturezas de operações são utilizadas no cálculo do superávit realista:

1- Operações financeiras pouco usuais envolvendo empresas que o governo controla, mas que estão fora do conceito de “setor público” para fins de apuração do

¹⁰ O período acima foi selecionado em razão da disponibilidade da série “resultado primário do setor público consolidado”, iniciada a partir de novembro de 2002, no site do Banco Central do Brasil (BCB).

¹¹ No caso Brasileiro, a estrutura organizacional do setor público propicia ações que visam aumentar o superávit primário por haver empresas (financeiras e não financeiras) nas quais o governo tem poder de mando, e que estão fora do conceito de “setor público” para fins de apuração do resultado primário e da dívida, como é o caso da Petrobrás e Eletrobrás. Este fato permite que o governo crie relações financeiras com as empresas que estão fora do conceito de setor público de modo a expandir suas despesas e sua dívida bruta sem, contudo, afetar o resultado primário e a dívida líquida.

¹² Journard et al. (2008)

resultado primário e da dívida, como é o caso da Petrobrás ou do BNDES.¹³ Chamaremos o conjunto dessas operações de X_1 .

2- Receitas não recorrentes (atípicas), que afetam o resultado fiscal de determinado exercício financeiro sem, no entanto, refletir o esforço fiscal daquele ano. O objetivo da maioria dessas operações é a obtenção de um aumento temporário das receitas para melhorar o resultado fiscal. Esse procedimento, no entanto, não contribui para a sustentabilidade das finanças públicas intertemporalmente. Chamaremos o conjunto dessas operações de X_2 .

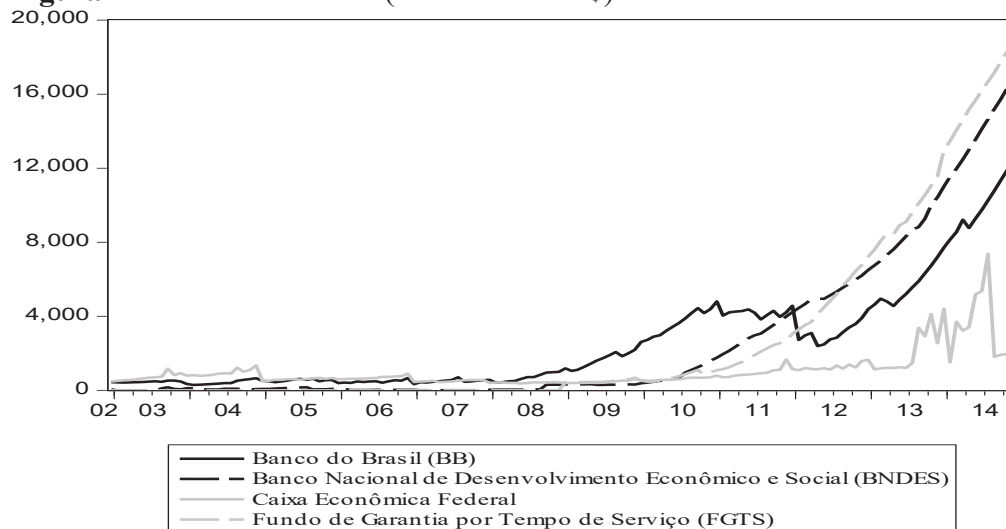
3- Atrasos nos repasses de verbas para os bancos públicos que financiam projetos sociais e previdenciários, fato que aumenta os passivos do Tesouro Nacional com as instituições públicas e infla artificialmente o superávit primário (“pedaladas fiscais”). Chamaremos o conjunto dessas operações de X_3 .

Assim, o superávit realista pode ser descrito como:

$$(2) \quad \text{real_PS} = \text{released_PS} - (X_1 + X_2 + X_3)$$

Na figura 2 abaixo, observa-se a crescente utilização das “pedaladas fiscais” a partir de 2009, devido à falta de recursos para financiar os programas sociais do governo, tais como o “Bolsa Família” e o “Minha Casa Minha Vida”. As pedaladas fiscais estão relacionadas à natureza das operações descritas no item 3 acima. Por sua vez, a Tabela 1 apresenta as operações do governo em termos de contabilidade criativa descritas nos itens 1 e 2 acima.¹⁴

Figura 2 – Pedaladas fiscais (em bilhões – R\$)



Fonte dos dados: Os dados foram obtidos do site do BCB – Nota à Imprensa de Política Fiscal – (<http://www.bcb.gov.br/?ECOIMPOLFISC>). Gráfico elaborado pelos autores. Nota: As instituições citadas são responsáveis por financiar programas sociais, previdenciários e de desenvolvimento. CEF: financia o programa “Bolsa Família”, o Seguro Desemprego, entre outros; FGTS: financia o programa “Minha Casa Minha Vida”; BB: financia as equalizações agrícolas; BNDES: financia o Programa de sustentação do investimento (PSI).

¹³ No caso Brasileiro, a estrutura organizacional do setor público propicia ações que visam aumentar o superávit primário por haver empresas (financeiras e não financeiras) nas quais o governo tem poder de mando, e que estão fora do conceito de “setor público” para fins de apuração do resultado primário e da dívida. Este fato permite que o governo crie relações financeiras com as empresas que estão fora do conceito de setor público de modo a expandir suas despesas e sua dívida bruta sem, contudo, afetar o resultado primário e a dívida líquida.

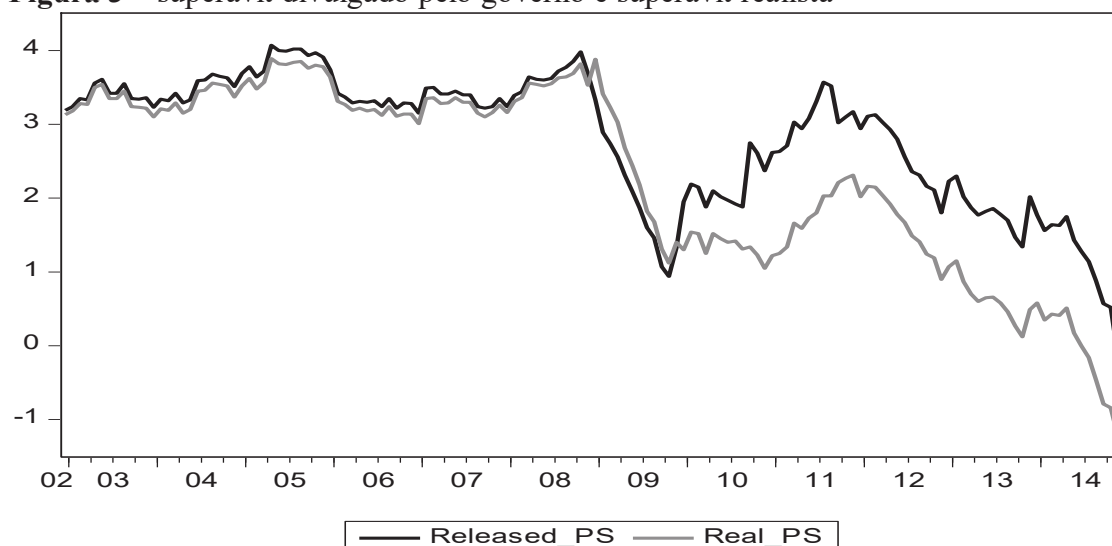
¹⁴ Os dados da Tabela 1 foram obtidos do “Anexo do Tesouro Nacional” (<https://www.tesouro.fazenda.gov.br/resultado-do-tesouro-nacional>), da Lei de Acesso à Informação (E-sic) e da Receita Federal (<http://www.receita.fazenda.gov.br/Historico/Arrecadacao/ResultadoArrec>).

Tabela 1 – Contabilidade criativa (em bilhões – R\$)

	Antecipação de dividendos	REFIS	Antecipação de receita do Pré-sal (Petrobrás)	Depósitos Judiciais	Contribuição sobre o Lucro Líquido (CSLL) da Vale	Lucros Sacados do Fundo Nacional de Desenvolvimento	Receitas e despesas no Fundo Soberano do Brasil	Bônus de concessão do Campo de Libra*	Total de contabilidade de criativa	Resultado Primário Oficial	Resultado Primário Realista	Meta superávit primário (% PIB)
2003	1,049								0.1%	3.2%	3.1%	4.3%
2004	1,367					396			0.1%	3.7%	3.6%	4.5%
2005	565					456			0.0%	3.7%	3.7%	4.3%
2006	1,011					1,234			0.1%	3.2%	3.1%	4.3%
2007	613					581			0.0%	3.2%	3.2%	3.8%
2008	1,644					2,06	-22,123		-0.6%	3.3%	3.9%	3.8%
2009	595	4,738		8,900		4,264			0.6%	1.9%	1.3%	2,5%
2010	4,478	7,689	31,879	4,000					1.3%	2.6%	1.3%	3.1%
2011	3,396	21,019			5,800				0.7%	2.9%	2.2%	3.1%
2012	7,598	16,559					15,466		0.9%	2.2%	1.3%	3.1%
2013	6,447	13,933						15,000	0.7%	1.7%	1.0%	2,3%
2014	6,392	9,371							0.3%	0.5%	0.2%	1,9%

Tabela elaborada pelos autores. Nota: REFIS: plano de refinanciamento de dívidas tributárias lançado no período da crise de 2009. *O governo antecipou o leilão para o fim de 2013, para que o pagamento fosse realizado até o fim do ano.

A variável *transparency* varia entre zero e um. Quando se aproxima de zero, isto significa uma maior opacidade no resultado fiscal, ou seja, os números apresentados pelo governo são resultados de contabilidade criativa para melhorar seus indicadores. Por sua vez, quando se aproxima de um isto quer dizer que menos artifícios contábeis foram utilizados e, portanto, maior a transparência na divulgação dos dados relativos a real situação fiscal. Espera-se uma correlação positiva da transparência fiscal com a credibilidade fiscal. A figura 3 apresenta o superávit divulgado pelo governo (*released_PS*, em preto), e o superávit realista (*real_PS*, em cinza), e evidencia, a partir de 2010 em diante, a recorrente interferência do governo por meio de artifícios contábeis para inflar o resultado primário.

Figura 3 – superávit divulgado pelo governo e superávit realista

Fonte: Gráfico elaborado pelos autores.

Outra variável, que representa o compromisso do governo em atingir as metas de superávit primário, é chamada *fiscal_commitment*. A variável é formada pela razão

entre o superávit primário divulgado e a meta: $fiscal_commitment = \frac{released_PS}{PS_target}$.¹⁵ Se o governo atinge sucessivamente a meta, o público irá esperar comprometimento do governo com seus objetivos, e, assim, a credibilidade fiscal será melhorada. Logo, quanto maior for esta variável, maior é o comprometimento e o sucesso do governo com as metas de superávit e, portanto, maior a credibilidade fiscal. Considerando-se ainda o impacto da solvência fiscal sobre a credibilidade, é utilizada uma variável referente ao espaço fiscal, nomeada *fiscal_space*. A variável representa quantos anos o governo levaria para pagar a dívida baseado nas receitas obtidas, ou seja, a variável expressa o tamanho da dívida frente à capacidade de pagá-la, e quanto maior seu valor, menos espaço fiscal dispõe o governo. Assim, $fiscal_space = \frac{debt_gdp}{revenue_gdp}$.¹⁶ É esperada uma correlação negativa entre *fiscal_space* e a credibilidade fiscal.

A análise também utiliza a dívida pública líquida como percentual do PIB (*debt_gdp*), a qual quando aumenta piora a solvência fiscal e afeta diretamente a expectativa sobre a dívida e a credibilidade fiscal, e; a taxa de juros implícita (*implicit_interest*), a qual reflete o gasto efetivo com os passivos e ativos da dívida pública.¹⁷ A tabela 2 apresenta as estatísticas descritivas das séries.

Tabela 2 – Estatísticas descritivas

Desc. Statistics	Variables						
	<i>fc</i>	<i>afc</i>	<i>debt_gdp</i>	<i>implicit_interest</i>	<i>fiscal_commitment</i>	<i>transparency</i>	<i>fiscal_space</i>
Mean	0.454	0.405	42.117	15.832	0.766	0.785	2.633
Median	0.491	0.410	40.995	15.638	0.800	0.957	2.451
Maximum	1.000	0.810	60.110	19.075	1.130	1.224	4.200
Minimum	0.026	0.020	31.050	13.461	-0.090	0.002	1.849
Std. Dev.	0.239	0.167	7.661	1.188	0.186	0.284	0.611
Skewness	0.213	0.160	0.366	0.423	-1.222	-0.823	0.717
Kurtosis	2.546	3.156	2.128	2.706	5.668	2.688	2.582
Observations	144	144	144	144	144	144	144

Fonte: Tabela elaborada pelos autores.

A tabela 3 apresenta as correlações entre as séries e a tabela 4 os testes de causalidade de Granger¹⁸. Em relação às principais variáveis de interesse sendo analisadas (*transparency* e *fiscal_space*), é possível observar correlação positiva entre a transparência e a credibilidade, e que a direção de causalidade é da transparência para a credibilidade. Por sua vez, é observada correlação negativa entre *fiscal_space* e a credibilidade, e que a direção de causalidade é da variável *fiscal_space* para a credibilidade. As demais variáveis também apresentaram as correlações esperadas.

¹⁵ O superávit primário divulgado é o “Resultado primário total do setor público consolidado (% PIB)” (acumulado em 12 meses), obtido do BCB - SGST (código 5793).

¹⁶ Para exemplo de aplicações desta variável, ver Aizenman et al.(2010) e Peat et al. (2015). A variável “*revenue_gdp*” são as “Receitas do Tesouro Nacional (Fluxos) (acumulado em 12)”, obtida do BCB - SGST (código 2266), como percentual do PIB (PIB em valores correntes, obtida do BCB - SGST (código 4380).

¹⁷ A variável *debt_gdp* é a “Dívida líquida do setor público total (DLSP) (% PIB) - Setor público consolidado”, obtida do BCB - SGST (código 4513). A variável *implicit_interest* é a Série histórica da “taxa de juros implícita (DLSP), acumulada em 12 meses”, obtida do BCB - Tabelas Especiais (www.bcb.gov.br/TAXADLSP).

¹⁸ Pelo fato do índice de credibilidade ser construído com base nas expectativas e essas poderem se alterar rapidamente, o teste de causalidade de Granger foi realizado com apenas uma defasagem.

Tabela 3 – Correlações

<i>Variables</i>	<i>fc</i>	<i>afc</i>
<i>debt_gdp</i>	-0.36	-0.16
<i>fiscal_space</i>	-0.43	-0.23
<i>implicit_interest</i>	-0.48	-0.39
<i>fiscal_commitment</i>	0.03	0.08
<i>transparency</i>	0.27	0.42

Fonte: Tabela elaborada pelos autores.

Tabela 4 – Causalidade de Granger

Null Hypothesis:	Obs	F-Statistic	Prob.
<i>transparency</i> does not Granger Cause <i>fc</i> <i>fc</i> does not Granger Cause <i>transparency</i>	143	15.8231 0.14287	0.0001 0.706
<i>fiscal_commitment</i> does not Granger Cause <i>fc</i> <i>fc</i> does not Granger Cause <i>fiscal_commitment</i>	143	35.4902 0.1004	0.0000 0.7517
<i>fiscal_space</i> does not Granger Cause <i>fc</i> <i>fc</i> does not Granger Cause <i>fiscal_space</i>	143	5.3201 1.6417	0.0225 0.2022
<i>debt_gdp</i> does not Granger Cause <i>fc</i> <i>fc</i> does not Granger Cause <i>debt_gdp</i>	143	5.5302 0.7373	0.0201 0.3920
<i>implicit_interest</i> does not Granger Cause <i>fc</i> <i>fc</i> does not Granger Cause <i>implicit_interest</i>	143	0.8380 7.5865	0.3615 0.0067
<i>transparency</i> does not Granger Cause <i>afc</i> <i>afc</i> does not Granger Cause <i>transparency</i>	143	15.2996 0.05601	0.0001 0.8133
<i>fiscal_space</i> does not Granger Cause <i>afc</i> <i>afc</i> does not Granger Cause <i>fiscal_space</i>	143	5.4144 2.75797	0.0214 0.099
<i>fiscal_commitment</i> does not Granger Cause <i>afc</i> <i>afc</i> does not Granger Cause <i>fiscal_commitment</i>	143	35.5749 0.24194	0.0000 0.6236
<i>debt_gdp</i> does not Granger Cause <i>afc</i> <i>afc</i> does not Granger Cause <i>debt_gdp</i>	143	5.1661 1.6110	0.0246 0.2065
<i>implicit_interest</i> does not Granger Cause <i>afc</i> <i>afc</i> does not Granger Cause <i>implicit_interest</i>	143	1.5014 9.8637	0.2225 0.0021

Fonte: Tabela elaborada pelos autores.

2.2 Metodologia

Pelo fato da variável dependente (credibilidade fiscal) ser uma variável contínua no intervalo entre zero e um, as estimações foram feitas utilizando o modelo TOBIT (Tobin, 1958). Como salientado por McDonald e Moffitt (1980), a aplicação do método é adequado e conveniente quando se assume que a variável dependente tem seus valores agrupados dentro de um limite. O modelo TOBIT usa todas as informações, incluindo informações acerca do censuramento, e fornece estimativas consistentes dos parâmetros (Long, 1997). Assim, a opção de truncamento entre zero e um foi utilizada, e esses valores foram imputados na opção de estimação por TOBIT do software E-Views9. As especificações para a análise por meio do modelo TOBIT são:

- (3) $fc = \beta_0 + \beta_1 debt_gdp + \beta_2 implicit_interest + \varepsilon_{1t}$
- (4) $fc = \alpha_0 + \alpha_1 debt_gdp + \alpha_2 implicit_interest + \alpha_3 fiscal_commitment + \varepsilon_{2t}$
- (5) $fc = \gamma_0 + \gamma_1 debt_gdp + \gamma_2 implicit_interest + \gamma_3 transparency + \varepsilon_{3t}$
- (6) $fc = \delta_0 + \delta_1 debt_gdp + \delta_2 implicit_interest + \delta_3 fiscal_commitment * transparency + \varepsilon_{5t}$
- (7) $fc = \beta_3 + \beta_4 fiscal_space + \beta_5 implicit_interest + \varepsilon_{6t}$
- (8) $fc = \alpha_4 + \alpha_5 fiscal_space + \alpha_6 implicit_interest + \alpha_7 fiscal_commitment + \varepsilon_{7t}$
- (9) $fc = \gamma_4 + \gamma_5 fiscal_space + \gamma_6 implicit_interest + \gamma_7 transparency + \varepsilon_{8t}$
- (10) $fc = \delta_4 + \delta_5 fiscal_space + \delta_6 implicit_interest + \delta_7 fiscal_commitment * transparency + \varepsilon_{10t}$

Em que ε_i representa o termo de erro aleatório.

O conjunto de especificações acima foi definido para verificar: (i) se o grau de interferência do governo nos resultados para o superávit primário por meio de artifícios contábeis e contabilidade criativa (medido pela variável *transparency*) afeta a credibilidade fiscal; (ii) se o esforço do governo em cumprir as metas de superávits contribui para o ganho de credibilidade fiscal, e; (iii) se o *fiscal_space* afeta a credibilidade fiscal. As estimações foram repetidas para todas as equações utilizando *afc* como variável dependente.

2.3 Resultados

A Tabela 5 apresenta as estimações para o conjunto de equações em que *fc* é a variável dependente e a Tabela 6 apresenta as estimações com a variável dependente *afc*.¹⁹

Em relação aos coeficientes estimados das variáveis *debt_gdp* e *implicit_interest*, todos apresentaram os sinais esperados e significância estatística. Assim, um aumento na variável *debt_gdp* impacta negativamente a credibilidade fiscal. Por sua vez, um aumento na taxa de juros implícita impacta negativamente a credibilidade fiscal, uma vez que aumenta o valor dos juros pagos pela dívida pública.

Em relação às variáveis de maior interesse, as estimações revelam que quando o governo atinge as metas anunciadas de superávit primário (*fiscal_commitment*), a confiança do público aumenta em relação ao comprometimento do governo e, assim, a credibilidade fiscal aumenta. Ou seja, quanto mais próximo da meta, ou acima dela está o superávit primário, maior é o impacto positivo sobre a credibilidade fiscal. Em relação à variável *transparency*, os resultados apontam que o público percebe a interferência do governo nos resultados relacionados ao superávit primário por meio da contabilidade criativa. Portanto, a transparência fiscal é um elemento importante para a credibilidade fiscal. Os achados sugerem que interferências do governo nas contas públicas por meio de contabilidade criativa com intuito de inflar artificialmente o superávit primário são percebidas negativamente pelo público e acarretam perda de confiança no governo. Nesse sentido, resultados criados por meio de artifícios contábeis e contabilidade criativa em relação ao superávit primário reduzem a transparência fiscal e levam a deterioração da credibilidade fiscal.

No tocante aos achados para a variável *fiscal_space*, os coeficientes estimados indicam que a solvência fiscal é determinante para a construção da credibilidade fiscal. Assim, um aumento na variável *fiscal_space* (a qual indica quantos anos o governo levaria para pagar a dívida baseado nas receitas obtidas) gera um impacto negativo sobre a credibilidade fiscal, devido à maior dificuldade em manter a trajetória da dívida sustentável.

De modo a responder de que forma o público reage aos descumprimentos da meta de superávit primário em um contexto no qual o governo é mais transparente em termos fiscais, um termo de interação (*transparency X fiscal_commitment*) foi testado. Os resultados sugerem que em um ambiente com transparência, os agentes confiam mais nos resultados divulgados em relação ao superávit primário e, assim, dão maior peso ao cumprimento da meta de superávit primário, pois, $\delta_3 > \alpha_3$ e $\delta_7 > \alpha_7$

¹⁹ Pelo fato de não haver concentração de valores da variável dependente nas situações limites (0 e 1), as estimações também foram realizadas por meio do método de mínimos quadrados ordinários, e as estimativas para os coeficientes permanecem as mesmas daquelas obtidas por TOBIT.

Os resultados são corroborados quando utilizamos a variável dependente *afc* (Tabela 6). Ou seja, os sinais e a significância estatística são observados para as estimativas com a variável dependente *afc*.

Tabela 5 - Estimações TOBIT (variável dependente: *fc*)

Dependent Variable: <i>fc</i>	TOBIT							
REGRESSORS	Eq3	Eq4	Eq5	Eq6	Eq7	Eq8	Eq9	Eq10
C	2.206*** (0.230) [9.561]	2.119*** (0.259) [8.158]	1.659*** (0.180) [9.209]	2.197*** (0.233) [9.408]	2.096*** (0.238) [8.782]	1.975*** (0.268) [7.356]	1.289*** (0.194) [6.641]	1.897*** (0.192) [9.866]
<i>debt_gdp</i>	-0.008*** (0.002) [-3.556]	-0.010*** (0.002) [-4.661]	-0.031*** (0.002) [-12.002]	-0.022*** (0.002) [-7.984]				
<i>fiscal_space</i>					-0.130*** (0.029) [-4.345]	-0.154*** (0.026) [-5.832]	-0.380*** (0.031) [-12.269]	-0.286*** (0.032) [-8.834]
<i>implicit_interest</i>	-0.087*** (0.016) [-5.213]	-0.088*** (0.016) [-5.350]	-0.032** (0.014) [-2.247]	-0.072*** (0.015) [-4.617]	-0.082*** (0.017) [-4.746]	-0.082*** (0.016) [-4.897]	-0.026* (0.014) [-1.828]	-0.063*** (0.012) [-5.068]
<i>fiscal_commitment</i>		0.227** (0.106) [2.139]				0.244** (0.105) [2.307]		
<i>transparency</i>			0.819*** (0.063) [12.846]				0.745*** (0.056) [13.232]	
<i>transparency X commitment</i>				0.532*** (0.088) [6.043]				0.517*** (0.072) [7.158]
Left Censored Obs	0	0	0	0	0	0	0	0
Right Censored Obs	3	3	3	3	3	3	3	3
Uncensored Obs	141	141	141	141	141	141	141	141

Fonte: Tabela elaborada pelos autores. Nota: Níveis de significância: ***denota 1%, **denota 5% e *denota 10%. Erro-padrão entre parênteses e estatística t entre colchetes. Todas as estimativas por TOBIT utilizaram a abordagem de covariância robusta de Huber-White.

Tabela 6 - Estimações TOBIT (variável dependente: *afc*)

Dependent Variable: <i>afc</i>	TOBIT							
REGRESSORS	Eq3	Eq4	Eq5	Eq6	Eq7	Eq8	Eq9	Eq10
C	1.316*** (0.179) [7.336]	1.263*** (0.196) [6.438]	0.916*** (0.140) [6.501]	1.311*** (0.181) [7.226]	1.298*** (0.187) [6.939]	1.222*** (0.204) [5.988]	0.701*** (0.153) [4.569]	1.161*** 0.182 6.363
<i>debt_gdp</i>	-0.001 (0.001) [-0.995]	-0.003* (0.001) [-1.725]	-0.018*** (0.002) [-9.137]	-0.011*** (0.002) [-5.023]				
<i>fiscal_space</i>					-0.037 (0.023) [-1.579]	-0.053** (0.021) [-2.487]	-0.223*** (0.023) [-9.323]	-0.146*** 0.024 -5.916
<i>implicit_interest</i>	-0.052*** (0.013) [-4.014]	-0.053*** (0.012) [-4.142]	-0.012 (0.011) [-1.071]	-0.042*** (0.012) [-3.481]	-0.050*** (0.013) [-3.695]	-0.050*** (0.013) [-3.828]	-0.008 (0.011) [-0.781]	-0.037*** 0.012 -3.098
<i>fiscal_commitment</i>		0.142 (0.087) [1.620]				0.155* (0.087) [1.770]		
<i>transparency</i>			0.596*** (0.049) [12.152]				0.551*** (0.044) [12.410]	
<i>transparency X commitment</i>				0.363*** (0.071) [5.106]				0.360*** 0.065 5.492
Left Censored Obs	0	0	0	0	0	0	0	0
Right Censored Obs	0	0	0	0	0	0	0	0
Uncensored Obs	144	144	144	144	144	144	144	144

Fonte: Tabela elaborada pelos autores. Nota: Níveis de significância: ***denota 1%, **denota 5% e *denota 10%. Erro-padrão entre parênteses e estatística t entre colchetes. Todas as estimativas por TOBIT utilizaram a abordagem de covariância robusta de Huber-White.

3. Análise de robustez

De modo a dar robustez aos resultados encontrados e verificar o efeito das principais variáveis analisadas considerando diferentes níveis de credibilidade fiscal,

realizamos estimações por regressão quantílica. Introduzido por Koenker e Basset (1978), o método de regressão quantílica divide a distribuição de uma maneira que uma dada proporção de observações esteja localizada abaixo do quantil.²⁰ Neste sentido, é possível observar o coeficiente estimado para diferentes partes da distribuição da credibilidade fiscal. O método de regressão quantílica permite observar os efeitos dos regressores para diferentes níveis de credibilidade fiscal.

Assim, as especificações 4, 5, 6, 8, 9 e 10 serão estimadas por regressão quantílica de modo a captar o efeito das principais variáveis analisadas no estudo. A tabela 7 apresenta as estimações por regressão quantílica para as especificações escolhidas em que a variável dependente é *fc*, e a figura 4 ilustra os resultados por meio de gráficos para o comportamento dos coeficientes de cada um dos regressores em cada especificação.

Com base nas estimações, é possível observar que o efeito de *debt_gdp* e o efeito de *fiscal_space* sobre a credibilidade fiscal diminuem na medida em que aumenta a credibilidade fiscal. Ou seja, os coeficientes estimados das duas variáveis são negativos, e na medida em que a credibilidade fiscal aumenta, o valor dos coeficientes estimados diminui em magnitude. O aumento da dívida pública como proporção do PIB (o qual também é captado pela variável *fiscal_space*) deteriora ainda mais a credibilidade fiscal quando esta se encontra em níveis mais baixos. Por sua vez, na medida em que aumenta a credibilidade fiscal, a magnitude do coeficiente estimado para a taxa de juros implícita também aumenta.

Os achados sugerem que o compromisso com a meta de superávit primário é fundamental para o governo conquistar credibilidade fiscal, principalmente quando o governo possui baixa credibilidade. Os resultados apontam que *fiscal_commitment* tem maior impacto sobre a credibilidade fiscal quando esta última está em níveis mais baixos, e o efeito dessa variável diminui na medida em que a credibilidade fiscal aumenta. Nesse sentido, quando o governo atua de maneira comprometida com seu objetivo de cumprir a meta de superávit primário durante a fase em que ainda está construindo sua credibilidade (e, portanto, a credibilidade fiscal ainda é baixa), o efeito desse comprometimento sobre a construção de credibilidade fiscal é maior. Sendo assim, se o governo deseja acelerar o processo de conquista de credibilidade fiscal quando esta ainda está em construção, é fundamental que atue de maneira comprometida com meta de superávit primário. Por outro lado, quando o governo possui níveis elevados de credibilidade fiscal, o efeito da falta de compromisso sobre a credibilidade perde significância.

Em relação ao efeito da transparência fiscal, os achados revelam que sua importância é maior para a credibilidade fiscal na fase em que a credibilidade está mais baixa. Isto quer dizer também que quando a credibilidade fiscal é baixa e o governo atua por meio de contabilidade criativa de modo a inflar artificialmente o superávit primário – o que aumenta a opacidade fiscal – o impacto adverso sobre a credibilidade fiscal é ainda maior. Portanto, se o governo se encontra em situação de baixa credibilidade fiscal, mas deseja tomar medidas que aumentem a credibilidade, uma delas é não fazer uso de contabilidade criativa, pois com isso a transparência fiscal aumentará (e em contrapartida a opacidade fiscal irá diminuir). De todo modo, mesmo quando a credibilidade se situa em um patamar maior, os efeitos da falta de transparência não podem ser considerados desprezíveis, apesar de serem inferiores a quando a credibilidade é baixa.

²⁰ Nós estimamos a regressão quantílica com *moving blocks bootstrap* (MBB), o qual fornece erros-padrão robustos para heteroscedasticidade e autocorrelação de forma desconhecida (Fitzenberger, 1997).

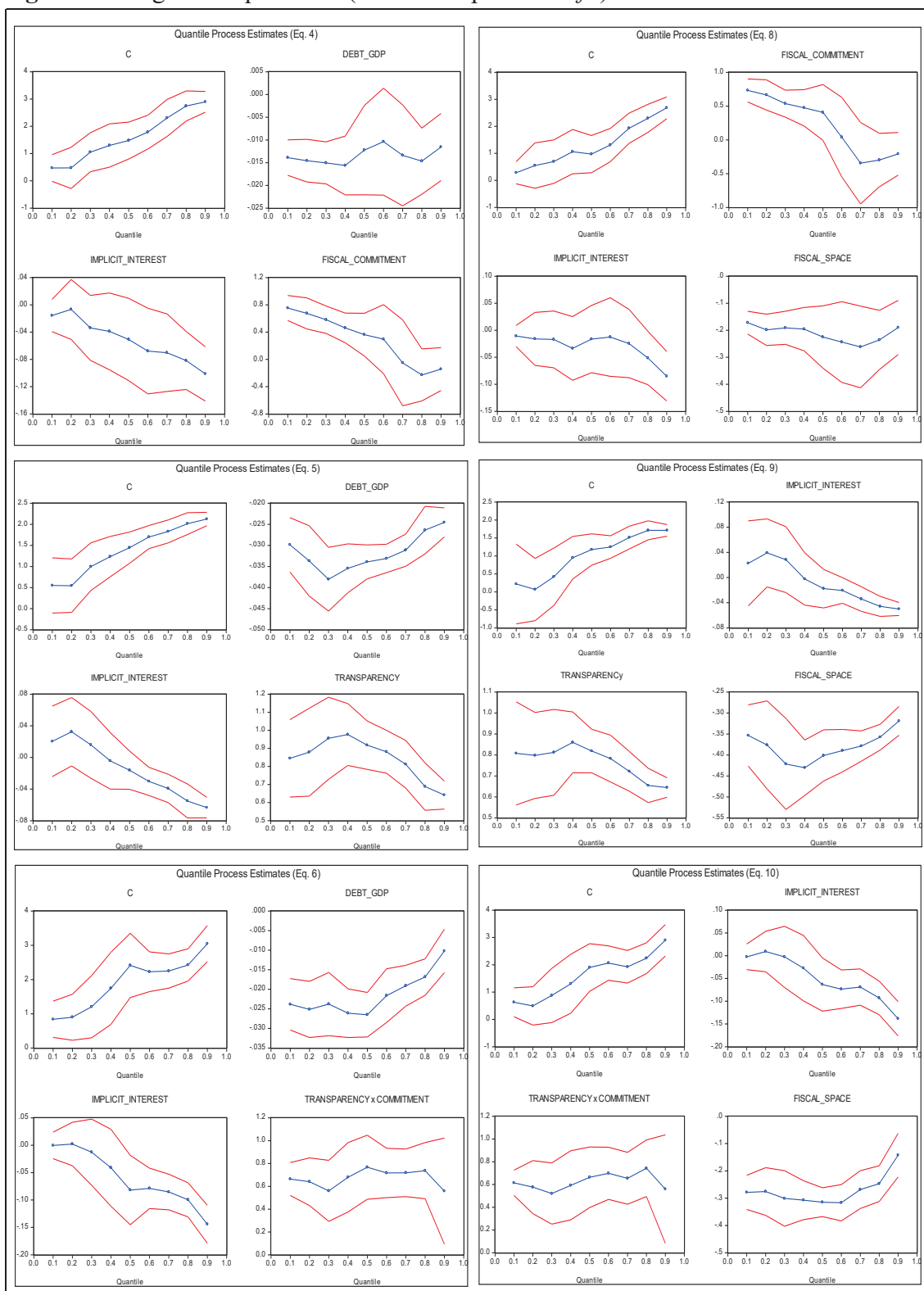
As mesmas estimações foram realizadas para as especificações em que a variável dependente é *afc*. De um modo geral, os resultados permanecem os mesmos.

Tabela 7 - Regressão quantílica (variável dependente: *fc*)

Specifications		Eq. 4		Eq. 5		Eq. 6		Eq. 8		Eq. 9		Eq. 10	
Regressors	Quantile	Coef.	Prob.	Coef.	Prob.	Coef.	Prob.	Coef.	Prob.	Coef.	Prob.	Coef.	Prob.
<i>constant</i>	0.1	0.4703	0.0263	0.5504	0.1018	0.8388	0.0023	0.2821	0.2098	0.2210	0.5445	0.6297	0.0212
	0.2	0.4758	0.2096	0.5434	0.0944	0.8996	0.0096	0.5484	0.1426	0.0674	0.8568	0.4937	0.1700
	0.3	1.0496	0.0024	0.9956	0.0007	1.2068	0.0104	0.6945	0.0815	0.4219	0.2518	0.8775	0.0854
	0.4	1.2970	0.0010	1.2336	0.0000	1.7425	0.0015	1.0634	0.0155	0.9508	0.0003	1.3054	0.0192
	0.5	1.4777	0.0000	1.4454	0.0000	2.4114	0.0000	0.9722	0.0110	1.1792	0.0000	1.9013	0.0000
	0.6	1.7856	0.0000	1.6969	0.0000	2.2269	0.0000	1.3080	0.0002	1.2485	0.0000	2.0633	0.0000
	0.7	2.3057	0.0000	1.8309	0.0000	2.2484	0.0000	1.9296	0.0000	1.5118	0.0000	1.9261	0.0000
	0.8	2.7428	0.0000	2.0159	0.0000	2.4255	0.0000	2.2957	0.0000	1.7140	0.0000	2.2400	0.0000
	0.9	2.8962	0.0000	2.1246	0.0000	3.0459	0.0000	2.6845	0.0000	1.7121	0.0000	2.8959	0.0000
<i>debt_gdp</i>	0.1	-0.0139	0.0000	-0.0299	0.0000	-0.0238	0.0000						
	0.2	-0.0146	0.0000	-0.0337	0.0000	-0.0251	0.0000						
	0.3	-0.0151	0.0000	-0.0380	0.0000	-0.0238	0.0000						
	0.4	-0.0156	0.0000	-0.0354	0.0000	-0.0261	0.0000						
	0.5	-0.0122	0.0410	-0.0339	0.0000	-0.0265	0.0000						
	0.6	-0.0104	0.1615	-0.0331	0.0000	-0.0216	0.0000						
	0.7	-0.0134	0.0541	-0.0311	0.0000	-0.0191	0.0000						
	0.8	-0.0147	0.0024	-0.0264	0.0000	-0.0169	0.0000						
	0.9	-0.0116	0.0073	-0.0245	0.0000	-0.0102	0.0004						
<i>implicit_interest</i>	0.1	-0.0157	0.1360	0.0203	0.3738	-0.0004	0.9754	-0.0107	0.3410	0.0225	0.3074	-0.0025	0.8614
	0.2	-0.0067	0.7618	0.0324	0.1439	0.0021	0.9158	-0.0160	0.4750	0.0392	0.0800	0.0091	0.6877
	0.3	-0.0338	0.1682	0.0158	0.4635	-0.0129	0.6772	-0.0169	0.5619	0.0284	0.2443	-0.0026	0.9387
	0.4	-0.0389	0.1864	-0.0043	0.8128	-0.0411	0.2528	-0.0337	0.3225	-0.0022	0.9027	-0.0275	0.4552
	0.5	-0.0508	0.1479	-0.0162	0.1906	-0.0819	0.0123	-0.0164	0.6472	-0.0176	0.1872	-0.0634	0.0353
	0.6	-0.0678	0.0754	-0.0303	0.0011	-0.0788	0.0000	-0.0125	0.7596	-0.0205	0.0231	-0.0736	0.0008
	0.7	-0.0702	0.0411	-0.0393	0.0000	-0.0854	0.0000	-0.0246	0.4669	-0.0341	0.0004	-0.0689	0.0009
	0.8	-0.0816	0.0044	-0.0551	0.0000	-0.0997	0.0000	-0.0513	0.0576	-0.0458	0.0000	-0.0928	0.0000
	0.9	-0.1013	0.0000	-0.0635	0.0000	-0.1442	0.0000	-0.0851	0.0006	-0.0500	0.0000	-0.1385	0.0000
<i>fiscal_commitment</i>	0.1	0.7529	0.0000					0.7318	0.0000				
	0.2	0.6742	0.0000					0.6645	0.0000				
	0.3	0.5823	0.0000					0.5356	0.0000				
	0.4	0.4618	0.0028					0.4739	0.0113				
	0.5	0.3625	0.0814					0.4061	0.1161				
	0.6	0.2961	0.3262					0.0400	0.9146				
	0.7	-0.0506	0.8849					-0.3451	0.3421				
	0.8	-0.2276	0.3071					-0.2987	0.1815				
	0.9	-0.1429	0.4337					-0.2074	0.2493				
<i>transparency</i>	0.1			0.8447	0.0000					0.8075	0.0000		
	0.2			0.8789	0.0000					0.7977	0.0000		
	0.3			0.9557	0.0000					0.8122	0.0000		
	0.4			0.9768	0.0000					0.8598	0.0000		
	0.5			0.9179	0.0000					0.8187	0.0000		
	0.6			0.8811	0.0000					0.7827	0.0000		
	0.7			0.8119	0.0000					0.7218	0.0000		
	0.8			0.6886	0.0000					0.6546	0.0000		
	0.9			0.6416	0.0000					0.6447	0.0000		
<i>transparency x commitment</i>	0.1					0.6639	0.000					0.6149	0.0000
	0.2					0.6400	0.000					0.5772	0.0000
	0.3					0.5602	0.000					0.5212	0.0002
	0.4					0.6790	0.000					0.5930	0.0002
	0.5					0.7671	0.000					0.6633	0.0000
	0.6					0.7169	0.000					0.6982	0.0000
	0.7					0.7181	0.000					0.6547	0.0000
	0.8					0.7370	0.000					0.7427	0.0000
	0.9					0.5580	0.020					0.5599	0.0229
<i>fiscal_space</i>	0.1							-0.1719	0.0000	-0.3536	0.0000	-0.2787	0.0000
	0.2							-0.1985	0.0000	-0.3763	0.0000	-0.2755	0.0000
	0.3							-0.1911	0.0000	-0.4216	0.0000	-0.3010	0.0000
	0.4							-0.1957	0.0000	-0.4304	0.0000	-0.3074	0.0000
	0.5							-0.2251	0.0003	-0.4012	0.0000	-0.3147	0.0000
	0.6							-0.2436	0.0044	-0.3898	0.0000	-0.3168	0.0000
	0.7							-0.2618	0.0006	-0.3789	0.0000	-0.2685	0.0000
	0.8							-0.2360	0.0000	-0.3576	0.0000	-0.2468	0.0000
	0.9							-0.1899	0.0008	-0.3188	0.0000	-0.1425	0.0006

Fonte: Tabela elaborada pelos autores.

Figura 4 – Regressão quantílica (variável dependente: fc)



Fonte: Gráfico elaborado pelos autores.

Conclusão

Este estudo mostrou evidências empíricas de que a falta de transparência derivada de contabilidade criativa causa danos à credibilidade fiscal. Ademais, os achados

apontam que o compromisso com a estabilidade fiscal é uma conduta que deve ser seguida por governos que desejam guiar as expectativas dos agentes e melhorar a confiança do público em relação à condução da política fiscal e ao gerenciamento da dívida pública.

Uma contribuição deste trabalho foi reformular o índice de credibilidade fiscal proposto por de Mendonça e Machado (2013) de modo a apurar a atual situação fiscal do Brasil e melhor compreender seus determinantes. O indicador revela que, desde 2009, a credibilidade fiscal está se deteriorando.

O trabalho também contribui com a elaboração de um indicador de transparência fiscal, o qual captura o grau de interferência do governo no resultado primário devido à contabilidade criativa. Os resultados apontam que práticas contábeis pouco usuais são percebidas e mal vistas pelo público, e repercutem de forma contrária à pretendida por governos que desejam inflar os resultados fiscais. Além disso, analisamos o quanto o público leva em conta o compromisso do governo com o lado fiscal em um ambiente com maior transparência. A partir das estimações, observamos que o público torna-se mais exigente quando o governo é mais transparente, penalizando mais o governo quando este descumpre as metas de superávit, uma vez que confia mais nos resultados divulgados. Nesse sentido, ações do governo que reduzem a transparência (e aumentam a opacidade) desestimulam uma postura de maior compromisso fiscal.

Estimações realizadas por regressão quantílica apontam importantes achados no tocante à mudança na magnitude dos efeitos das principais variáveis analisadas, quando considerados diferentes níveis de credibilidade fiscal. Na medida em que o governo apresenta credibilidade mais elevada, o tamanho da dívida pública (ou da variável *fiscal space*) exerce um efeito negativo cada vez menor sobre a credibilidade. Em relação à transparência fiscal, os resultados mostram que aumentos na transparência causam um impacto maior na credibilidade quando esta última está em níveis mais baixos. Já a falta de compromisso com a meta de superávit torna-se cada vez menos nociva para a credibilidade fiscal quando a credibilidade chega a patamares mais elevados. Esses achados sugerem que quando o governo atinge um nível de credibilidade suficientemente elevado, este possui maior grau de liberdade para conduzir a política econômica. Nesse contexto, o governo seria capaz de fazer uso de políticas anticíclicas quando necessário, uma vez que a maior credibilidade torna o governo menos suscetível a avaliações negativas por parte do público. Fica, portanto, a reflexão de que ações sucessivas que revelam falta de compromisso fiscal tornam o governo temporalmente inconsistente e podem gerar reversão das expectativas do público em relação à sustentabilidade da dívida pública e alcance das metas de superávit, levando a novos ciclos de deterioração da credibilidade. Sugere-se, assim, graduar uma eventual expansão fiscal com o compromisso contínuo com a estabilidade da dívida.

Referências

- Aizenman, J., Jinjark, Y., 2010. De Facto Fiscal Space and Fiscal Stimulus. NBER No. 16539.
- Aizenman, J., Hutchison, M., Jinjark, Y., 2013. What is the risk of European sovereign debt defaults? Fiscal space, CDS spreads and market pricing of risk. *Journal of International Money and Finance*, 34, 37-59.
- Alt, J. E. , Lassen, D. D., 2006. Transparency, political polarization, and political budget cycles in OECD countries. *American Journal of Political Science*, 50(3).
- Arbatli, E., Escolano, J., 2012. Fiscal Transparency, Fiscal Performance and Credit Ratings. IMF Working Paper.

- Bernoth, K., Wolff, G. B., 2008. Fool the markets? Creative accounting, fiscal transparency and sovereign risk premia. *Scottish Journal of Political Economy*, 55(4), 465-487
- Blanchard, O., Cottarelli, C., 2010. Ten commandments for fiscal adjustment in advanced economies. The International Monetary Fund's Global Economy Forum, July.
- Bunch, B. S., 1991. The effect of constitutional debt limits on state governments' use of public authorities. *Public Choice*, 68, 57-69.
- Carranza L., Daude C., Melguizo A., 2014. Public infrastructure investment and fiscal sustainability in Latin America: incompatible goals? *Journal of Economic Studies*, 41(1), 29-50.
- Celasun, O., Debrun, X., Ostry, J. D., 2006. Primary Surplus Behavior and Risks to Fiscal Sustainability in Emerging Market Countries: A "Fan-Chart" Approach. *IMF Staff Papers*, 53(3), 401-425
- de Mendonça, H. F., Machado, M. R., 2013. Public Debt Management and Credibility: Evidence from an emerging economy. *Economic Modelling*, 30(1), 10-21.
- de Mendonça, H. F., Tostes, I., 2015. The Effect of Monetary and Fiscal Credibility on Exchange Rate Pass-Through in an Emerging Economy. *Open Economies Review*, 26(4), 787-816
- Fitzenberger, B., 1997. A Guide to Censored Quantile Regressions. In: *Handbook of Statistics, Robust Inference* (Eds. G.S. Maddala & C.R. Rao). 15, 405-437.
- FMI, 2002. Assessing sustainability. Policy Development and Review Department, May. Disponível em: <https://www.imf.org/external/np/pdr/sus/2002/eng/052802.pdf>
- FMI, 2012 Fiscal Transparency, Accountability, and Risk, FMI Policy Paper (Washington), August. Disponível em: <http://www.imf.org/external/np/pp/eng/2012/080712.pdf>
- Glennerster, R., & Shin, Y., 2008. Does Transparency Pay? IMF staff papers, 55(1).
- Journard, I., Minegishi, M., André, C., Chantal Nicq, C., Price, R., 2008. Accounting for one-off operations when assessing underlying fiscal positions. OECD Economics Department Working Papers, n. 642. Paris: OECD.
- Kiewiet, D. and Szakaly, K., 1996. Constitutional limitations on borrowing: analysis of state bonded indebtedness. *Journal of Economics, Law, and Organization*, 62-97.
- Koen, V. and Van Den Noord, P., 2005. Fiscal Gimmickry in Europe: One-Off Measures and Creative Accounting, OECD Economics Department Working Paper, 417, OECD Publishing.
- Koenker, R., Bassett, G., 1978. Regression Quantiles. *Econometrica*. 46(1), 33-50.
- Kopits, G., Cragg, J., 1998. Transparency in government operations. IMF Occasional paper.
- Levy, E., 2010. Fiscal (and External) Sustainability. MPRA Paper No. 23956, July.
- Long, J. S., 1997. Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage Publications
- McDonald, J.F. and Moffitt, A., 1980. The uses of Tobit analysis. *The review of economic and statistics*, 62, (2), 318-321.
- Milesi-Ferretti, G., 2004. Good, bad or ugly? On the effects of fiscal rules with creative accounting. *Journal of Public Economics*, 88 (1), 377-394.
- Montes, G. C., Acar, T., 2015 Determinants of Fiscal Credibility and Country Risk in Brazil: An Empirical Analysis. *The Empirical Economics Letters*, 14, 38-43.
- Peat, M., Svec, J., Wang, J., 2015. The effects of fiscal opacity on sovereign credit spreads. *Emerging Markets Review*, 24, 34-45.
- Petersen, J., 2003. Changing red to black: deficit closing alchemy. *National Tax Journal*. 56 (3).
- Schettini, B. P., Gouvêa, R. R., Orair, R. O., Gobetti, S. W. Resultado Fiscal estrutural e impulso fiscal: uma aplicação para as administrações públicas no Brasil, 1997-2010. Brasília: IPEA No. 1650
- Ter-Minassian, T., 2010. Preconditions for a successful introduction of structural fiscal balance-based rules in Latin America and the Caribbean: a framework. Washington: BID. (Discussion Paper, No. 157).
- Tobin J., 1958. Estimation of relationships for limited dependent variables, *Econometrica* 26, 24-36.

- Von Hagen, J. and Wolff, G. B., 2006. What do deficits tell us about debt? Empirical evidence on creative accounting with fiscal rules in the EU. *Journal of Banking and Finance*, 30 (12) 3259–79.
- Von Hagen, J., 1991. A note on the empirical effectiveness of formal fiscal restraints. *Journal of Public Economics*, 44, 199–210.

Apêndice

Tabela A.1 Alterações na meta de superávit primário (em % do PIB)

	LDO	Meta	Alteração na LDO	Nova Meta	Outras alterações
2009	Lei nº 11.768	3,8%	Lei nº 12.053 (09/10/2009)	2,5%	Empresas do Grupo Petrobras não serão consideradas no cálculo de superávit primário. -Amplia o limite de abatimento no superávit devido ao Programa de Aceleracao do Crescimento (PAC) de R\$ 15,5 bilhões para R\$ 28,5 bilhões.
2010	Lei nº 12.017	3,3%	Lei nº 12.377 (30/12/2010)	3,10%	- Eletrobrás também não será considerada no cálculo do superávit - Amplia o limite de abatimento no superávit devido ao Programa de Aceleracao do Crescimento (PAC) de R\$ 22,5 bilhões para R\$ 29,8 bilhões
2011	Lei nº 12.309	3,1%	-	-	-
2012	Lei nº 12.465	3,1%	-	-	-
2013	Lei nº 12.708	3,1%	Lei nº 12.901 (18/12/2013)	2,3%	- Amplia o limite de abatimento no superávit devido ao Programa de Aceleracao do Crescimento (PAC) de R\$ 45,2 bilhões para R\$ 65,2 bilhões
2014	Lei nº 12.919	1,9%	PL nº 36/2014 (11/11/2014)	-	- Acaba com o limite fixo de abatimento de R\$ 67 bilhões do PAC e dá liberdade para abater a totalidade dos gastos do PAC, além da inclusão das desonerações tributárias.

Fonte: Tabela elaborada pelos autores. Nota: LDO Lei de Diretrizes Orçamentárias; PL Projeto de Lei

Prevendo a inflação no Brasil com grande conjunto de dados: uma aplicação do modelo de fatores comuns

Cristiano Santos^{}, Roberto Tatiwa Ferreira[†], Ivan Castelar[‡]*

25 de abril de 2016

Resumo

Este artigo compara as previsões da inflação acumulada em 12 meses fornecidas por modelos índice de difusão que usam fatores comuns estimados de um grande conjunto de dados, com as previsões fornecidas pela pesquisa Focus do Banco Central do Brasil e por modelos autor-regressivo integrado de média móvel (ARIMA) e vetor autorregressivo (VAR). Os resultados mostram que as previsões da pesquisa Focus são as que apresentam menor erro quadrático médio entre os modelos comparados e que os ganhos de se utilizar grande quantidade de dados para prever a inflação são limitados. Apenas o modelo índice de difusão que usa pré-seleção de variáveis na estimação dos fatores comuns adiciona informações à previsão da pesquisa Focus, de forma que o erro quadrático médio da previsão combinada é menor do que o erro quadrático médio de cada previsão isoladamente.

Palavras-chave: Previsão. Inflação. Modelo de fatores. Grande conjunto de dados.

Abstract

This paper compares the forecasts of inflation in 12 months provided by models diffusion index using a large data set, by Focus survey of Brazil's Central Bank, by integrated autoregressive models moving average (ARIMA) and by vector autoregression (VAR). The results show that the Focus survey forecasts are those with lower mean square error between the models compared and that the gains of using large data set to predict inflation are limited. Only diffusion index forecasts using targeted predictors add information to forecast the Focus survey, so that the mean square error of the combined forecast is lower than the mean square error of each forecast alone.

keywords: Forecast. Inflation. Factor Models. Large Data Sets.

JEL Classification: E31, E37.

Área 4: Macroeconomia Aplicada (Applied Macroeconomics)

^{*}Universidade Federal de Alagoas-UFAL. E-mail: cristiano1br@gmail.com

[†]Universidade Federal do Ceará-CAEN/UFC. E-mail: rtf2@uol.com.br

[‡]Universidade Federal do Ceará-CAEN/UFC. E-mail: lume1250@yahoo.com.br

1 Introdução

Dado que as mudanças na política monetária têm efeitos defasados sobre a economia, um dos principais objetivos ao se prever a inflação é auxiliar a autoridade monetária na definição de seus instrumentos de política no presente e no futuro próximo. No caso do Brasil, em que o Banco Central (BC) utiliza um sistema de metas, é mais relevante prever se a inflação acumulada nos próximos 12 meses estará dentro da meta do que prever a inflação em um mês específico.

Considerando a importância do problema de prever a inflação e a crescente quantidade de séries macroeconômicas disponíveis aos tomadores de decisão, este artigo prevê a inflação acumulada em 12 meses por modelos índice de difusão que usam um grande conjunto de dados e compara com as previsões fornecidas pela pesquisa Focus do BC do Brasil e por modelos autorregressivo integrado de média móvel (ARIMA) e vetor autorregressivo (VAR).

O modelo de índice de difusão (DI) proposto por Stock e Watson (2002) resume as informações contidas num grande número de séries temporais em poucos fatores comuns, que são então inseridos numa regressão simples para prever a inflação. Duas configurações diferentes deste mesmo modelo propostas por Bai e Ng (2008) e Dias et al. (2010) são também usadas neste artigo. A ideia básica é unir as metodologias do modelo Índice de Difusão com preditores e fatores *targeted*. Apesar de simples, não encontramos na literatura nenhum trabalho que utilize essa extensão do modelo DI. Assim, este artigo também contribui ao propor esse novo modelo e ao investigar sua aplicabilidade na previsão da inflação no Brasil.

Os resultados mostram que as previsões da pesquisa Focus são as que apresentam menor erro quadrático médio entre os modelos comparados, o que indica a importância da previsão da Focus para monitorar a política monetária. O desempenho dos modelos índice de difusão aponta que os ganhos de se utilizar grande quantidade de dados para prever a inflação se mostraram limitados. Apenas o modelo índice de difusão com pré-seleção de preditores de Bai e Ng (2008) adiciona informações à previsão da pesquisa Focus, de forma que o erro quadrático médio (EQM) da previsão combinada é menor do que o EQM de cada previsão isoladamente.

O trabalho está organizado da seguinte forma: A seção 2 fornece uma breve revisão da literatura. A seção 3 diz respeito ao modelo de fatores comuns e procedimento de estimação. A seção 4 descreve os modelos índice de difusão e demais modelos usados nas previsões. A seção 5 apresenta a metodologia de avaliação e os dados utilizados. A seção 6 demonstra os resultados e discussões. Por fim, a seção 7 apresenta as conclusões.

2 Revisão da literatura

Previsões usando fatores comuns extraídos de um grande conjunto de séries temporais ficaram conhecidas na literatura como previsão com índice de difusão, devido aos trabalhos de Stock e Watson (1998, 2002), mostrando a eficiência da metodologia. A partir de então, vários autores tem investigado a aplicação de modelos de fatores comuns para previsão das principais variáveis macroeconômicas, como a inflação e o produto, comparando com outros modelos tradicionais de previsão.

Em um trabalho extenso, Artis et al. (2005) mostram evidências que o uso dos fatores comuns na regressão de previsão melhora as previsões para as variáveis reais e de preço do Reino Unido, quando comparadas com modelos AR e VAR de séries temporais.

Usando dados da África do Sul, Gupta e Kabundi (2011) encontram que as previsões do modelo de fatores comuns para a inflação, crescimento do produto e taxa de juros, apresentam erros

menores que as previsões dos modelos VAR, VAR Bayesiano e Equilíbrio Geral Dinâmico Estocástico (DSGE). Para a Austrália, o trabalho de Moser et al. (2007) encontra que a inflação prevista com modelo de fatores supera as previsões dos modelos ARIMA e VAR em termos de eficiência preditiva.

Gavin e Kliesen (2008) utilizam modelos de fatores para prever a inflação e o produto nos Estados Unidos e descobrem que esses modelos apresentam previsões significantemente melhores que os modelos univariados. Eles ainda destacam que esses modelos são mais úteis quando o horizonte de previsão é longo. Com resultados semelhantes, Cheung e Demers (2007) indicam que o modelo índice de difusão fornece consideráveis melhorias de precisão na previsão do PIB e do núcleo da inflação do Canadá quando comparado ao modelo autorregressivo, principalmente quando o horizonte de previsão aumenta. Eles também revelam que não há diferença em termos de previsão ao usar fatores estáticos ou dinâmicos generalizados.

Schumacher e Dreger (2004) indicam que a previsão da taxa de crescimento do PIB da Alemanha com modelos de fatores comuns tem erros menores que os modelos alternativos. No entanto, essa redução não é estatisticamente significativa e concluem que o ganho de usar esse tipo de modelo na amostra por eles analisada é limitada. Com dados da área do euro, Angelini et al. (2001) reportam que não existem ganhos significativos ao usar modelos de fatores como previsor da inflação neste grupo de países.

No caso do Brasil, ainda há poucos trabalhos aplicados. Ferreira et al. (2005) usam modelos índice de difusão para prever a taxa de crescimento do PIB trimestral e encontram que esses modelos tem desempenho melhor que modelos autorregressivos. Com uma abordagem semelhante a este artigo, Figueiredo (2010) mostra que modelos de fatores podem gerar melhores previsões para a taxa de inflação brasileira, principalmente em horizontes de previsão mais longos. No entanto, Figueiredo (2010) prevê apenas a inflação mensal, enquanto neste artigo a previsão é para inflação acumulada em 12 meses utilizando novas metodologias de previsão com modelos de fatores.

Para uma revisão completa sobre modelos de fatores e previsões com índice de difusão ver o trabalho de Eickmeier e Ziegler (2008). Nele, os autores fazem uma meta-análise da grande literatura existente e descobrem que modelos de fatores possuem desempenho de previsão melhor para as variáveis dos Estados Unidos do que para Reino Unido, que o tamanho da base de dados afeta positivamente o desempenho e que pré-selecionar as variáveis incluídas na base de dados não têm influência significativa.

3 Modelo de Fatores Comuns

No modelo de fatores comuns, cada variável é representada como a soma de dois componentes não observados: o componente comum e o componente idiossincrático. O componente comum é formado por um pequeno número de fatores comuns que afetam todas as variáveis enquanto que o componente idiossincrático representa choques específicos a cada variável.

Dependo da forma como os fatores comuns influenciam as variáveis, o modelo de fatores é classificado como estático, se os fatores afetam as variáveis apenas contemporaneamente, ou dinâmico, se fatores também afetam as variáveis defasadamente, ou seja, se a influência dos fatores persiste por alguns períodos.

3.1 Modelo de Fatores Comuns Estático

Seja $X_t = (X_{1t}, X_{2t}, \dots, X_{Nt})'$, para $t = 1, \dots, T$, um vetor $N \times 1$ de séries temporais estacionárias com média zero. O modelo de fatores comuns supõe que cada variável X_{it} admite a seguinte representação:

$$X_{it} = \lambda_i' F_t + e_{it} \quad (1)$$

onde F_t é um vetor ($r \times 1$) de fatores comuns, λ_i é um vetor ($r \times 1$) de fatores *loadings* e e_{it} é o componente idiossincrático.

Nesse modelo há r fatores comuns F_t que sumarizam a estrutura de covariância de todas as variáveis X_{it} . A forma e a intensidade da influência de um F_t sobre uma variável específica i dependem de λ_i . Em outras palavras, a ideia do modelo é que a dimensão dos fatores é bem menor que a dimensão dos dados, $r \ll N$, de forma que um pequeno número de fatores é capaz de replicar grande parte da variância observada nas variáveis X_{it} .

Considerando a série temporal multivariada X_t de dimensão ($N \times 1$), a equação (1) pode ser reescrita como:

$$X_t = \Lambda F_t + e_t \quad (2)$$

onde $\Lambda_{(N \times r)} = (\lambda_1, \lambda_2, \dots, \lambda_r)'$ e $e_{t(N \times 1)} = (e_{1t}, e_{2t}, \dots, e_{Nt})'$

Usando notação matricial:

$$X = F\Lambda' + e \quad (3)$$

onde $X = (X_1, X_2, \dots, X_N)$ é uma matriz $T \times N$ com os dados observados, $F = (F_1, \dots, F_T)'$ é uma matriz $T \times r$ de fatores comuns e $e = (e_1, \dots, e_T)'$ é uma matriz $T \times N$ de componentes idiossincráticos.

Bai e Ng (2002) e Bai (2003) mostram que os pesos (*loadings*) Λ dos fatores podem ser estimados pelo método de componentes principais da seguinte forma:

$$\hat{\Lambda} = \sqrt{N}V \quad (4)$$

$$\hat{F} = X\hat{\Lambda}/N \quad (5)$$

onde V é uma matriz $N \times r$ com as colunas sendo os r maiores auto vetores da matriz $X'X$. A estimativa da matriz residual é obtida por $\hat{e} = X - \hat{F}\hat{\Lambda}'$.

É importante destacar que $\hat{F}$ estima FH (uma rotação de F) e $\hat{\Lambda}$ estima ΛH^{-1} (uma rotação de Λ), onde H é qualquer matriz $r \times r$ invertível. Desta forma não é possível identificar os fatores comuns e *loadings* de forma única, embora o produto $\hat{F}\hat{\Lambda}$ estime $F\Lambda$ diretamente. Esta falta de identificação não é problemática para previsão, mas deve ser considerada quando os fatores são interpretados de maneira estrutural.

Para que os fatores possam ser estimados consistentemente pelo método de componentes principais, Bai e Ng (2002) e Bai (2003) mostram que os seguintes pressupostos devem ser atendidos¹:

- *Pressuposto A – Fatores comuns:* $E\|F_t\|^4 < \infty$ e $T^{-1} \sum_{t=1}^T F_t F_t' \xrightarrow{p} \Sigma_F$ quando $T \rightarrow \infty$ para alguma matriz $r \times r$ positiva definida Σ_F
- *Pressuposto B – Fatores loadings:* $\|\lambda_i\| \leq \bar{\lambda} < \infty$ e $\|\frac{\Lambda'\Lambda}{N} - \Sigma_N\| \rightarrow 0$ quando $N \rightarrow \infty$ para alguma matriz $r \times r$ positiva definida Σ_N .
- *Pressuposto C – Heterocedasticidade e Dependência no tempo e no cross-section:* Existe uma constante positiva $M < \infty$, tal que para todo N e T ,

¹Os pressupostos A-D são os mesmos do trabalho de Bai e Ng (2002)

1. $E(e_{it}) = 0$ e $E|e_{it}|^8 < \infty$
2. $E(e'_s e_t / N) = E(N^{-1} \sum_{i=1}^N e_{is} e_{it}) = \gamma_N(s, t)$, $|\gamma_N(s, t)| \leq M$ para todo s e

$$T^{-1} \sum_{s=1}^T \sum_{t=1}^T |\gamma_N(s, t)| \leq M$$

3. $E(e_{it} e_{jt}) = \tau_{ij,t}$ com $|\tau_{ij,t}| < |\tau_{ij}|$ para algum τ_{ij} e para todo t . Em adio,

$$N^{-1} \sum_{i=1}^N \sum_{j=1}^N |\tau_{ij}| \leq M$$

4. $E(e_{it} e_{js}) = \tau_{ij,ts}$ e $(NT)^{-1} \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^T \sum_{s=1}^T |\tau_{ij,ts}| \leq M$
5. Para todo (t, s) , $E|N^{-1/2} \sum_{i=1}^N [e_{is} e_{it} - E(e_{is} e_{it})]|^4 \leq M$

- Pressuposto D – Dependncia fraca entre fatores e erros idiossincrticos:

$$E \left(\frac{1}{N} \sum_{i=1}^N \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^T F_t e_{it} \right\| \right)$$

Os pressupostos A e B asseguram que os fatores comuns e loadings no so degenerados. O pressuposto C permite que os componentes idiossincrticos e_{it} possuam formas fracas de heterosedasticidade e sejam fracamente correlacionados no tempo e entre as sries. O pressuposto D est implcito nos pressupostos A e C e permite que os fatores comuns sejam fracamente correlacionados com os erros idiossincrticos.

3.2 Modelo de Fatores Comuns Dinmico

Uma generalizao do modelo de fatores comuns esttico em (1)  permitir que os fatores comuns afetem as sries de maneira contempornea e tambm defasada. Esta nova configurao  chamada modelo de fatores comuns dinmicos e pode ser expressa como:

$$X_t = \Lambda_0 \tilde{F}_t + \Lambda_1 \tilde{F}_{t-1} + \cdots + \Lambda_s \tilde{F}_{t-s} + e_t \quad (6)$$

onde $\tilde{F}_{t-j}$ so vetores $q \times 1$ de fatores comuns e Λ_j so matrizes $N \times q$ de fatores *loadings*, com $j = 0, 1, \dots, s$.

Se o nmero de defasagens s for finito, ento, como mostra Stock e Watson (1998) e Bai e Ng (2007), o modelo de fatores comuns dinmico em (6)  equivalente ao modelo de fatores comuns esttico:

$$X_t = \Lambda F_t + e_t \quad (7)$$

com $\Lambda = (\Lambda_0, \Lambda_1, \dots, \Lambda_s)$ e $F_t = (\tilde{F}_t', \tilde{F}_{t-1}', \dots, \tilde{F}_{t-s}')'$

Isto mostra que um modelo com q fatores dinmicos e s defasagens pode ser represenatado por um modelo com $r = q(s+1)$ fatores estticos. Esta equivalncia entre os modelos permite aplicar todos os resultados tericos - como estimao, especificao e inferncia - do modelo de fatores esttico ao modelo de fatores dinmico na forma esttica.

Vale ressaltar que o modelo de fatores comuns dinmicos pode permitir que as defasagens s sejam infinitas, neste caso os procedimentos de estimao usam o conceito de componente principal dinmico para estimar os fatores no domnio da frequncia, como pode ser visto nos trabalhos de Forni et al. (2000) e Forni e Lippi (2011).

4 Previsão com fatores comuns

Uma das principais aplicações do modelo de fatores comuns é na previsão de séries macroeconômicas, como foi visto na revisão de literatura. Esta seção descreve o modelo de fatores de Stock e Watson (2002) e duas versões modificadas desse modelo propostas por Bai e Ng (2008) e Dias et al. (2010).

4.1 Modelo Índice de Difusão

Suponha que um grande conjunto de séries temporais estacionárias $X_t = (X_{1t}, X_{2t}, \dots, X_{Nt})'$ são observadas para $t = 1, \dots, T$. O problema do agente econômico é prever y_{T+h} , estando em T . Uma solução para este problema é utilizar o modelo de Índice de Difusão (DI) proposto por Stock e Watson (2002) que consiste em usar os fatores comuns na seguinte equação de previsão:

$$y_{t+h} = \mu + \alpha(L)y_t + \beta(L)'F_t + \varepsilon_{t+h} \quad (8)$$

com F_t sendo um vetor $r \times 1$ de fatores comuns satisfazendo:

$$X_{it} = \lambda_i'F_t + e_{it} \quad (9)$$

onde $\alpha(L)$ e $\beta(L)$ são polinômios de defasagem de ordem p e m , respectivamente, y_t é a variável a ser prevista e μ é uma constante.

As equações (8) e (9) mostram como o modelo de fatores permite incorporar na previsão a informação de um grande número de preditores (X_t) de uma maneira simples e parcimoniosa. O termo “índice de difusão” foi usado em Stock e Watson (2002) para designar os fatores comuns que representam os co-movimentos das variáveis econômicas. A ideia básica é estimar um pequeno número de fatores que expliquem a maior parte da variabilidade das variáveis observadas em , para depois usá-los como preditores.

O modelo DI é estimado em um procedimento de dois estágios; primeiro obtém-se os fatores comuns usando o estimador de componentes principais apresentado em (5). Com esses fatores estimados², a equação de previsão em (8) é obtida usando Mínimos Quadrados Ordinários (MQO). O número de fatores comuns r e as ordens de defasagem p e m são escolhidas pelo Critério de Informação Bayesiano (BIC), com $1 \leq r \leq 4$, $1 \leq m \leq 3$ e $1 \leq p \leq 6$, onde os valores máximos escolhidos são os mesmo de Stock e Watson (2002).

Uma das críticas ao modelo DI de Stock e Watson (2002) é não considerar a variável específica que está sendo prevista. A partir desta crítica, dois refinamentos são propostos na literatura. O primeiro, denominado Índice de Difusão com preditores *targeted* (DI-tp), estabelece que as variáveis usadas para estimar os fatores comuns devem ser selecionadas considerando a variável específica a ser prevista. O segundo, chamado de Índice de Difusão com fatores *targeted* (DI-tf), propõe que se deve estimar o fator comum considerando a variável a ser prevista.

4.2 Modelo Índice de Difusão com preditores *targeted*

Apesar do modelo Índice de Difusão (DI) permitir a inclusão de um grande número de variáveis, o trabalho de Boivin e Ng (2006) tem evidenciado que simplesmente acrescentar mais

²Dado que o interesse é apenas na previsão gerada pelos fatores e não sobre os fatores em si, o problema de identificação inerente ao modelo de fatores é irrelevante neste caso.

variáveis para extrair os fatores comuns não é adequado quando o objetivo é previsão e as variáveis incluídas não são relevantes. Boivin e Ng (2006) mostram através de resultados empíricos que com um conjunto menor de variáveis macroeconômicas obtêm-se previsões com desempenho igual ou superior do que as obtidas com um conjunto maior de variáveis.

Motivados por este fato, Bai e Ng (2008) propõem um refinamento do modelo DI que denominaremos modelo Índice de Difusão com preditores *targeted* (DI-tp). O modelo DI-tp estabelece que as variáveis (preditores) X_{it} usadas para estimar os fatores comuns devem ser escolhidas considerando a variável específica y_t a ser prevista.

Por exemplo, se o objetivo é prever a inflação, então o conjunto de séries temporais usado para estimar os fatores comuns deve ser formado por variáveis que possuam alguma informação sobre a inflação futura.

Para determinar quais variáveis X_{it} são significativas para comporem o conjunto de dados usado na previsão de y_t , Bai e Ng (2008) propõe o método *hard thresholding* descrito a seguir:

- a) Para cada $i = 1, \dots, n$, efetua-se a regressão de y_t sobre W_{t-h} e $X_{i,t-h}$. Neste artigo W_{t-h} inclui uma constante e quatro defasagens de y_t .
- b) Denote por t_i a estatística t do parâmetro de $X_{i,t-h}$. c) Selecione os preditores $X_{i,t}$ cuja $|t_i|$ excede o nível de significância α .
- c) Estime F_t a partir do conjunto de preditores selecionados e siga o mesmo procedimento padrão do modelo DI.

Além deste método *hard thresholding*, Bai e Ng (2008) propõem outros métodos de seleção denominados *soft thresholding* que aplicam algoritmos como *least absolute shrinkage selection operator* (LASSO) e *least angle regressions* (LARS). Este artigo utiliza apenas o *hard thresholding*, já que eles concluem que ambos os métodos são superiores em termos de erro de previsão quando comparados ao modelo DI padrão.

4.3 Modelo Índice de Difusão com fatores *targeted*

O modelo Índice de Difusão com fatores *targeted* (DI-tf) é outra configuração do modelo DI proposta por Dias et al. (2010) na qual a estimação do fator comum leva em consideração a variável y_t que está sendo prevista.

Para isto, considere a seguinte modificação, seja $X_{(h)}$ uma matriz $(T - h) \times N$ com os dados sem as últimas h observações. Como já visto na seção 3, os fatores podem ser estimados como:

$$\hat{\Lambda}_{(h)} = \sqrt{N}V \quad (10)$$

$$\hat{F}_{(h)} = X\hat{\Lambda}_{(h)}/N \quad (11)$$

onde agora V é uma matriz $N \times N$ com as colunas sendo todos os auto vetores da matriz $X'_{(h)}X_{(h)}$.

A abordagem de Dias et al. (2010) consiste em gerar um fator comum *targeted* como uma combinação linear de todos os fatores estimados por componentes principais e a variável y_t a ser prevista da seguinte forma:

$$\hat{F}_{(h)t}^* = \sum_{n=1}^N \frac{\omega_{(h)n}}{\sum_{i=1}^N \omega_{(h)i}} \hat{F}_{(h)t,n} \quad (12)$$

com

$$\omega_{(h),n} = \left(\frac{1}{T-h} \sum_{t=1}^{T-h} \hat{F}_{(h)t,n} y_{t+h} \right) \frac{\mu_{(h),n}}{\mu_{(h),1}} \quad (13)$$

onde $\hat{F}_{(h)t,n}$ é o elemento da linha t e da coluna n da matriz $\hat{F}_{(h)}$, $\mu_{(h),n}$ são os auto valores de $X'_{(h)}X_{(h)}$.

Estimado o fator *targeted* $\hat{F}_{(h)t}^*$, a previsão segue o mesmo procedimento do modelo DI padrão. A diferença que neste caso a equação de previsão possui apenas um fator comum, $\hat{F}_{(h)t}^*$, que considera a variável a ser prevista em sua estimação.

4.4 Modelo Índice de Difusão com preditores e fatores *targeted*

Tendo apresentado duas formas diferentes do modelo DI como forma de melhorar as previsões, uma extensão natural é juntar essas duas formas de refinamento do modelo DI. Assim, o modelo Índice de Difusão com preditores e fatores *targeted* (DI-tfp) que é proposto neste artigo aplica a seleção de preditores igual ao modelo DI-tp e estima o fator comum da mesma forma que o modelo DI-tf, sendo, portanto, a combinação deste dois modelos. Desta forma, modelo DI-tfp considera a variável a ser prevista tanto na seleção dos preditores como na estimação dos fatores.

Apesar de simples, não encontramos na literatura nenhum trabalho que utilize essa extensão do modelo DI. Assim, este artigo também contribui ao propor esse novo modelo e ao investigar sua aplicabilidade na previsão da inflação no Brasil.

4.5 Previsão de referência e previsões alternativas

Para comparar as previsões fora da amostra dos modelos índice de difusão é necessário um modelo de referência (*benchmark*). Este artigo escolhe a previsão da pesquisa Focus do Banco Central (BC) do Brasil como referência e duas previsões alternativas geradas pelos modelos autorregressivo integrado de média móvel (ARIMA) e vetor autorregressivo (VAR).

O modelo ARIMA é simples e frequentemente produz previsões melhores que outros modelos econométricos, como aponta Hamilton (1994, seção 4.8). Para a especificação e estimação usa-se o método automatizado descrito em Hyndman e Khandakar (2008).

O modelo VAR usa duas ou mais variáveis assumindo que as mesmas se relacionam simultaneamente para gerar previsões. De maneira semelhante à Arruda et al. (2011), a escolha das variáveis usadas no VAR foi obtida a partir das possíveis combinações do conjunto de variáveis PIB, taxa de juros Selic e agregado monetário M1.

A pesquisa Focus fornece a média e a mediana das previsões para a taxa de inflação acumulada em 12 meses para o horizonte de 12 meses à frente obtidas pelo BC com cerca de 100 participantes do mercado financeiro. Neste artigo utilizamos a mediana das previsões por apresentar menor EQM do que média no período em estudo. Como as previsões são coletadas diariamente pela Focus, usou-se a média dessas previsões diárias para representar a previsão mensal da Focus em determinado mês. Isto torna o conjunto de previsões da Focus compatível ao dos modelos que utilizam dados mensais.

5 Previsão fora da amostra

O tipo de previsão fora da amostra usado neste artigo é recursivo, i.e., todas as previsões são baseadas nos valores passados da série até a data em que a previsão é realizada, de forma que a especificação e os parâmetros do modelo são atualizados em cada período usando os dados do início da amostra até data corrente. Este tipo de previsão fora da amostra é também usado em Stock e Watson (2002), Bai e Ng (2008) e Dias et al. (2010).

O procedimento de previsão fora da amostra é uma forma de se obter previsões que se aproximam das previsões obtidas por agentes econômicos em situações reais. A vantagem deste procedimento é que permite quantificar a capacidade de adaptação do modelo a novos dados.

Outra vantagem da previsão fora da amostra é permitir comparar a previsões obtidas de modelos econométricos com previsões obtidas de outras formas ou quando não se conhece o modelo que gerou a previsão. No caso deste artigo, usam-se as previsões da pesquisa Focus e, portanto, não se tem o modelo econométrico gerador das previsões, já que as mesmas são recolhidas de um grande grupo de profissionais do mercado financeiro.

Vale ressaltar dois pontos. Primeiro, que este procedimento não é indicado quando o objetivo é avaliar modelos com base na teoria econômica, pois um modelo teórico válido pode gerar previsões menos acuradas do que um modelo de série temporal, conforme discute Clements e Demers (1998) e Shmueli (2010). No entanto, quando o objetivo é apenas previsão, o procedimento fora da amostra é uma forma intuitiva e factível de se avaliar previsões.

O segundo ponto é que testes para previsão fora da amostra não devem ser aplicados para comparação de modelos, visto que neste caso, os testes estatísticos tradicionais que usam toda a amostra, como o teste F , apresentam melhores resultados em termos de poder estatístico (Diebold, 2015).

Para complementar a análise de previsão fora da amostra, dois testes de previsão comumente usados na literatura são descritos a seguir.

5.1 Teste de igual acurácia de previsão

Ao comparar os erros de previsão de dois modelos é necessário identificar se a diferença observada é estatisticamente significativa ou não. Os testes usados para responder essa questão são chamados na literatura como testes de igual acurácia de previsão (*equal forecast accuracy*).

O teste DM modificado é baseado em Diebold e Mariano (1995) e Harvey et al. (1997) e permite verificar se a diferença de acurácia de previsão de dois modelos concorrentes é estatisticamente significativa. A estatística do teste DM inicialmente proposta por Diebold e Mariano (1995) é a seguinte:

$$S_1 = \frac{\bar{d}}{\sum_{\tau=-(h-1)}^{h-1} \frac{1}{T} \sum_{t=|\tau|+1}^T (d_t - \bar{d})(d_{t-|\tau|} - \bar{d})} \quad (14)$$

onde a série perda diferencial $d_t = e_{A,t}^2 - e_{B,t}^2$ é a diferença dos erros quadrados das previsões³ de dois modelos A e B, $\bar{d}$ é a média amostral, o denominador é o estimador do erro padrão de $\bar{d}$ robusto a heteroscedasticidade e autocorrelação, T é o tamanho da série d_t e h é o horizonte de previsão.

³A série perda diferencial pode ser definida para qualquer função perda, $d_t = g(e_{A,t}) - g(e_{B,t})$, neste artigo foi escolhida a função perda quadrática para comparar o erro quadrático médio.

A hipótese nula do teste DM é de que os dois modelos tem o mesmo desempenho de previsão, ou seja, a diferença média entre os erros de previsão dos modelos é zero, $E(\bar{d}) = 0$. Diebold e Mariano (1995) mostram que sob a hipótese nula e assumindo que a série d_t é estacionária, a distribuição assintótica da estatística S_1 é normal padrão.

Harvey et al. (1997) sugerem um teste DM modificado para melhorar o desempenho do teste DM que consiste em utilizar a seguinte correção:

$$S_1^* = \sqrt{\frac{T+1-2h+T^{-1}h(h-1)}{T}} S_1 \quad (15)$$

onde S_1 é a estatística original descrita em (14).

Além desta correção, deve-se comparar a estatística S_1^* com os valores críticos da distribuição t com $T-1$ graus de liberdade para melhorar a performance do teste DM modificado em pequenas amostras.

5.2 Teste de previsão incorporada

Em previsão, mais importante que identificar o modelo com o menor erro de previsão é descobrir se modelos concorrentes possuem diferentes informações para prever uma variável de interesse, no sentido de que a combinação da previsão dos modelos seja melhor do que as previsões individuais.

Como observado por Harvey et al. (1998), um modelo concorrente que tenha alguma informação que é ausente no modelo de referência, é útil mesmo que possua erro de previsão maior, pois quando combina-se as previsões o erro diminui.

Essa questão da combinação de previsões de modelos rivais pode ser testada com uma simples regressão:

$$y_t = \alpha + (1 - \lambda)\hat{y}_{A,t} + \lambda\hat{y}_{B,t} + \varepsilon_t \quad (16)$$

onde $\hat{y}_{A,t}$ e $\hat{y}_{B,t}$ são as previsões para y_t obtidas de dois modelos A e B quaisquer, α é uma constante e ε_t é o erro da regressão.

O parâmetro λ é estimado por Mínimos Quadrados Ordinários (MQO) e a inferência é realizada por um teste t com o procedimento de correção para heterocedasticidade e autocorrelação de Newey e West (1987). A hipótese nula ($\lambda = 0$) é que o modelo A incorpora o modelo B em termos de previsão, ou seja, o modelo A incorpora a informação contida na previsão do modelo B . A hipótese alternativa ($\lambda > 0$) é que o modelo B possui informações ausentes no modelo A .

5.3 Dados

Como o objetivo do artigo é avaliar o uso da grande quantidade de séries macroeconômicas para prever a inflação através de modelos índice de difusão, foi coletado um conjunto de dados suficientemente grande, de forma a obter séries temporais regulares e representativas da economia brasileira.

O conjunto de dados, o qual é detalhado no apêndice A, consiste de 122 séries temporais macroeconômicas de frequência mensal, cobrindo o período de 1996.1 a 2014.2 e oriundas do Ipeadata e Banco Central do Brasil. As séries são agrupadas em sete categorias:

- (1) Produção (19 séries);
- (2) Preços e indicadores monetários (16 séries);
- (3) Mercado externo (27 séries);
- (4) Finanças públicas (14 séries);
- (5) Consumo e vendas (15 séries);
- (6) Mercado de trabalho (13 séries);
- (7) Mercado financeiro (18 séries).

Para tornar o comportamento das séries estacionário, foi usado dois tipos de transformações. Na transformação log-diferença, o logaritmo natural da série é diferenciado em um período, o que representa uma aproximação da mudança mensal em termos percentuais. Na transformação primeira-diferença, as séries são apenas diferenciadas em um período.

Como é típico na literatura, as séries em valores correntes foram deflacionadas usando um índice de preços, nesse artigo foi escolhido o IPCA. A correção para valores discrepantes (*outliers*) seguiu o mesmo procedimento automático de Stock e Watson (2005)⁴.

6 Resultados Empíricos e Discussão

Nesta seção são analisados os resultados das previsões para inflação no Brasil usando os modelos índice de difusão na previsão da inflação, a pesquisa Focus e os modelos ARIMA e VAR. Todos os resultados foram obtidos com o programa estatístico R (R Core Team, 2016). Os códigos e os dados utilizados, assim como detalhes para replicação, estão disponíveis em <https://github.com/santosc/prevendo.inflacao.fatores.comuns>

6.1 Comparando as previsões

O Índice de Preços ao Consumidor Amplo (IPCA) foi a medida de inflação escolhida por ser o índice oficial adotado pelo Banco Central (BC) do Brasil no sistema de metas de inflação. Os exercícios de previsão incluem dados de janeiro de 1996 até fevereiro de 2014. Este intervalo representa o período de estabilidade econômica após o plano real e possui uma grande quantidade de séries macroeconômicas disponíveis.

O IPCA acumulado em 12 meses é a série adotada pelo BC para monitorar o comportamento da inflação. Na publicação mensal “relatório da inflação”, o BC analisa e utiliza essa série, que ao fim do ano deve estar dentro da meta estabelecida para inflação. Assim, não é de grande relevância prever que a inflação de um determinado mês será elevada, se a inflação em 12 meses ficar dentro meta. Além disso, grande parte dos agentes econômicos, quando fecham contratos e estabelecem seus preços para um determinado ano, também está mais interessada na inflação acumulada em 12 meses do que na inflação de um determinado mês.

Dado a importância da inflação acumulada nos próximos 12 meses, esse artigo avalia a capacidade dos modelos em prever a série do IPCA em 12 meses que é mostrada na Figura 1.

O horizonte de previsão é de 12 meses à frente. Esse é o mesmo horizonte de previsão usado na pesquisa Focus para prever a série IPCA acumulada em 12 meses. Além de possibilitar a comparação com a Focus, essa escolha do horizonte de previsão também é útil para o BC, dado o efeito defasado da política monetária.

⁴O ajuste para *outlier* corresponde a substituição da observação com desvio absoluto maior do que seis vezes o intervalo interquartil pelo valor mediano das cinco observações anteriores.

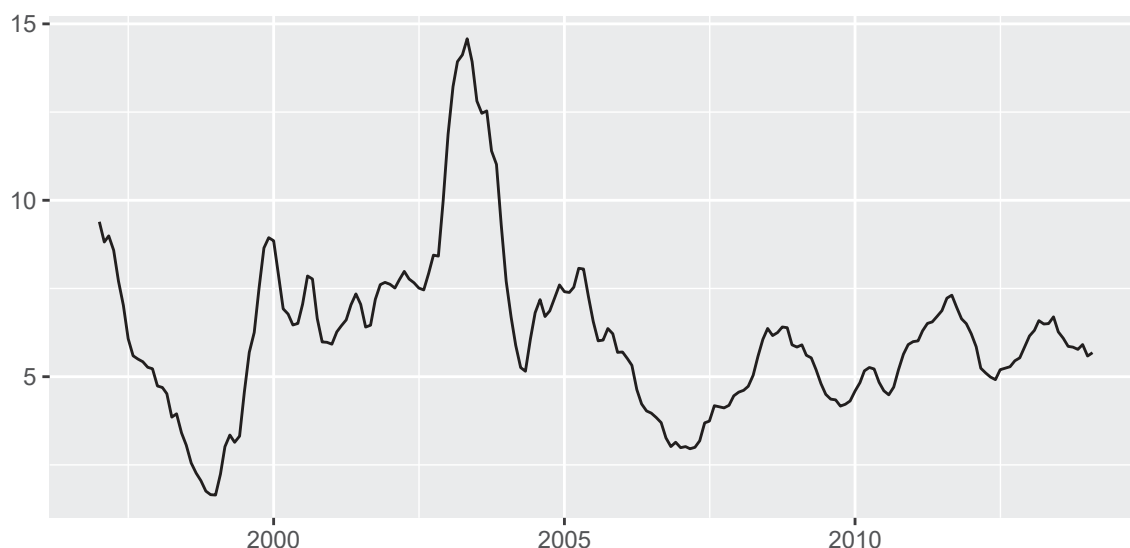


Figura 1: IPCA acumulado em 12 meses - 1996.1 - 2014.2

Fonte: Banco Central do Brasil. Elaboração dos autores.

Na simulação de previsão fora da amostra, todos os modelos foram estimados recursivamente do início da amostra até o período em que é realizada a previsão. As especificações dos modelos, como ordem de defasagem e número de fatores comuns, foram escolhidas pelo critério BIC.

Como medida da habilidade preditiva, utilizou-se a Raiz do Erro Quadrático Médio (REQM) e o Erro Quadrático Médio (EQM) relativo, que é simplesmente a razão entre o EQM do modelo e o EQM da pesquisa Focus. Um valor inferior a um desta razão indica de que o modelo possui um erro de previsão menor, ou seja, seu desempenho é melhor que a previsão da Focus.

A Tabela 1 apresenta a REQM e o EQM relativo dos modelos em relação à pesquisa Focus, junto com valor p do teste DM modificado para igualdade de previsão.

Tabela 1: Habilidade preditiva fora da amostra, 2008.5-2014.2

Modelo de Previsão	REQM ¹	EQM relativo ²	Teste DM ³
Pesquisa Focus	0,80	1,00	
ARIMA	1,02	1,64	0,33
VAR	1,58	3,92	0,00
Índice de Difusão (DI)	1,01	1,60	0,39
DI com fatores <i>targeted</i> (DI-tf)	1,03	1,67	0,38
DI com preditores <i>targeted</i> (DI-tp)	0,97	1,48	0,45
DI com fatores e preditores <i>targeted</i> (DI-tfp)	1,08	1,85	0,31

Fonte: Elaboração dos autores

¹ REQM: raiz do erro quadrático

² EQM relativo: EQM do modelo indicado dividido pelo EQM da pesquisa Focus.

³ valor p do teste DM modificado cuja hipótese nula é que os erros de previsão são estatisticamente iguais ao da pesquisa Focus.

A REQM de 0,80 para a pesquisa Focus indica que, ao comparar o valor do IPCA em 12 meses com o previsto pela Focus, comete-se um erro “típico” de 0,80% para mais ou menos. A mesma interpretação é dada para a REQM dos demais modelos.

O EQM relativo mostra que a previsão da Focus para a inflação acumulada possui erro quadrático médio inferior ao de todos os modelos avaliados no período de maio de 2008 a fevereiro de 2014. Isto indica que a Focus é um bom modelo de referência (*benchmark*) quando se prevê a inflação acumulada em 12 meses. Este apontamento é corroborado por Altug e Çakmakli (2016) que indicam que a pesquisa Focus para inflação acumulada também supera as previsões da curva de Phillips *forward* e *backward-looking*.

Considerando-se o teste DM modificado, não há diferença significativa entre as previsões da Focus e dos modelos econométricos estimados para o IPCA acumulado no período de avaliação considerado, com exceção do modelo VAR, cujo valor p foi menor que 1%. Assim, apesar da Focus possuir menor EQM, não foi encontrada diferença estatisticamente significativa em relação ao EQM dos demais modelos.

Considerando somente as previsões da Focus para inflação mensal sem as previsões para a inflação acumulada em 12 meses, Lima e Alves (2011, p.16) concluem que “não há qualquer evidência significativa de superioridade das previsões da Focus sobre modelos univariados simples em horizontes mais longos de previsão”. Esses resultados apontam que, para a inflação, a eficiência preditiva da Focus é diferente quando usa-se a série mensal ou série acumulada em 12 meses. Trabalhos futuros podem estudar as duas séries de previsões conjuntamente para confirmar e explicar essa diferença.

6.2 Comparando as previsões combinadas

Apesar da Focus possuir previsões com o menor EQM entre os modelos analisados, não se deve concluir que os modelos usados não são úteis para prever a inflação com base apenas na comparação dos erros. Outra questão importante é saber se algum dos modelos possui informação adicional que é ausente na previsão da Focus, de forma que a previsão combinada tenha erros menores.

A Tabela 2 apresenta o resultado do teste de previsão incorporada (*forecast encompassing*) de Harvey et al. (1998) que investiga se o modelo A (primeira coluna) incorpora as informações contidas nas previsões do modelo B (segunda coluna). Ou seja, o testa-se agora se a previsão de algum dos modelos contém informação que é ausente na previsão da Focus.

Tabela 2: Teste de previsão incorporada, 2008.5-2014.2

Modelo A	Modelo B	$\hat{\lambda}$ (valor p)
Pesquisa Focus	ARIMA	0,06 (0,85)
Pesquisa Focus	VAR	-0,17 (0,47)
Pesquisa Focus	Índice de Difusão (DI)	0,18 (0,26)
Pesquisa Focus	DI com fatores <i>targeted</i> (DI-tf)	-0,06 (0,83)
Pesquisa Focus	DI com preditores <i>targeted</i> (DI-tp)	0,24 (0,05)
Pesquisa Focus	DI com fatores e preditores <i>targeted</i> (DI-tfp)	0,03 (0,86)

Fonte: Elaboração dos autores

Nota: A tabela mostra o λ estimado na equação $y_t = \alpha + (1 - \lambda)\hat{y}_{A,t} + \lambda\hat{y}_{B,t} + \varepsilon_t$, onde y_t é o IPCA observado, $\hat{y}_{A,t}$ é a previsão da pesquisa Focus e $\hat{y}_{B,t}$ é a previsão do modelo na segunda coluna. A hipótese nula ($\lambda = 0$) é que o modelo B não adiciona poder preditivo ao modelo A .

O teste de previsão incorporada mostra que a previsão da Focus incorpora a previsão de todos os outros modelos, com exceção do modelo DI-tp. Ou seja, apenas o modelo DI-tp contém informação que não está contida nas previsões do mercado fornecidas pela Focus. Isto evidencia ganhos de informação ao selecionar os preditores antes da estimação dos fatores comuns, visto que o

mesmo não ocorreu para modelos DI e DI-tf.

Esta evidência corrobora os resultados de Boivin e Ng (2006) e Bai e Ng (2008) de que mais variáveis no conjunto de preditores nem sempre é mais adequado quando se usa o modelo de fatores para previsão. Desta forma, estudos sobre previsão usando modelo índice de difusão e fatores comuns devem considerar a forma de como o conjunto de preditores é selecionado em suas especificações.

Vale ressaltar que é difícil para um modelo simples, como DI-tp, adicionar informação às previsões da Focus, já que a mesma é uma combinação das previsões de vários especialistas do mercado financeiro que usam diversos métodos e possuem diferentes conjuntos de informação.

Para reforçar as evidências do teste de previsão incorporada, a Tabela 3 compara a habilidade preditiva da previsão da Focus combinada com cada um dos modelos por meio de uma média aritmética simples. As combinações apresentarem um EQM relativo menor que um, indicando melhora na previsão em termos de erro, com exceção da combinação com o modelo VAR.

Tabela 3: Habilidade preditiva fora da amostra das previsões combinadas, 2008.5-2014.2

Modelo de Previsão	REQM ¹	EQM relativo ²	Teste DM ³
Pesquisa Focus somente	0,80	1,00	
Focus e ARIMA	0,72	0,82	0,41
Focus e VAR	1,09	1,88	0,00
Focus e Índice de Difusão (DI)	0,70	0,77	0,06
Focus e DI com fatores <i>targeted</i> (DI-tf)	0,77	0,94	0,72
Focus e DI com preditores <i>targeted</i> (DI-tp)	0,68	0,73	0,04
Focus e DI com fatores e preditores <i>targeted</i> (DI-tfp)	0,73	0,85	0,51

Fonte: Elaboração dos autores

¹ REQM: raiz do erro quadrático

² EQM relativo: EQM do modelo indicado dividido pelo EQM da pesquisa Focus.

³ valor p do teste DM modificado cuja hipótese nula é que os erros de previsão são estatisticamente iguais ao da pesquisa Focus.

A previsão combinada com o modelo DI-tp apresentou uma redução no EQM de 27% em relação ao EQM da previsão somente da Focus, sendo a maior redução do EQM entre todas as combinações e a única com uma redução significativa ao nível de 5% conforme o teste DM modificado. Esta evidência reforça o resultado do teste de previsão incorporada da Tabela 2 mostrando que o modelo DI com preditores *targeted* apresenta informação que não consta nas previsões do mercado fornecidas pela pesquisa Focus.

Esta evidência também apoia o resultado encontrado por Figueiredo (2010) para inflação mensal no Brasil, de que o modelo de fatores com pré-seleção das variáveis (preditores *targeted*) apresenta melhor desempenho que outras abordagens de previsão com grande conjunto de dados, como Mínimos Quadrados Parciais (PLS).

7 Conclusão

Este artigo avaliou a aplicabilidade empírica de quatro modelos, índice de difusão (DI), índice de difusão com fatores *targeted* (DI-tf), índice de difusão com preditores *targeted* (DI-tp) e índice de difusão com preditores e fatores *targeted* (DI-tfp), que usam fatores comuns de um grande conjunto de séries macroeconômicas para prever a inflação no Brasil.

A avaliação se baseia em simulação de previsões fora da amostra, onde se compara os modelos índice de difusão com a pesquisa Focus do Banco Central do Brasil e com os modelos autorregressivo integrado de média móvel (ARIMA) e vetor autorregressivo (VAR).

As previsões da pesquisa Focus obtiveram a menor raiz do erro quadrático médio (REQM) entre todos os modelos quando se prevê a inflação acumulada um ano à frente. Isto ressalta a importância das previsões da Focus para monitorar a política monetária e para servir de referência (*benchmark*) quando se prevê a inflação acumulada em 12 meses.

Os resultados do teste de previsão incorporada indicam que apenas o modelo DI-tp adiciona informações à previsão da pesquisa Focus, de forma que a REQM da previsão combinada é menor do que a de cada previsão isoladamente. Nenhum outro modelo acrescentou informação às previsões da Focus de maneira significativa. Isto indica ganhos de informação ao selecionar os preditores de acordo com a variável a ser prevista. Vale ressaltar que estes resultados se limitam à inflação acumulada em 12 meses, podendo os modelos DI, DI-tf e DI-tpf ainda serem úteis para prever outras variáveis macroeconômicas.

Uma sugestão para trabalhos futuros é estudar modelos de fatores comuns para previsão que utilizem outros métodos de seleção dos preditores e verificar se as previsões da inflação melhoram em relação ao método de seleção usado neste artigo.

Agradecimentos

Os autores agradecem à Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (Funcap) pelo apoio financeiro.

Referências

- Altug, S. e C. Çakmakli (2016). Forecasting inflation using survey expectations and target inflation: Evidence for Brazil and Turkey. *International Journal of Forecasting* 32(1), 138–153.
- Angelini, E., J. Henry, e R. Mestre (2001). Diffusion Index-Based Inflation Forecasts for the Euro Area. *European Central Bank Working Paper April*(61), 1–47.
- Arruda, E. F., R. T. Ferreira, e I. Castelar (2011). Modelos Lineares e Não Lineares da Curva de Phillips para Previsão da Taxa de Inflação no Brasil. *Revista Brasileira de Economia* 65(3), 237–252.
- Artis, M. J., A. Banerjee, e M. Marcellino (2005). Factor forecasts for the UK. *Journal of Forecasting* 24(4), 279–298.
- Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions. *Econometrica* 71(1), 135–171.
- Bai, J. e S. Ng (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica* 70(1), 191–221.
- Bai, J. e S. Ng (2007). Determining the Number of Primitive Shocks in Factor Models. *Journal of Business & Economic Statistics* 25(1), 52–60.
- Bai, J. e S. Ng (2008). Large dimensional factor analysis. *Foundations and Trends in Econometrics* 3(2), 89–163.

- Boivin, J. e S. Ng (2006). Are more data always better for factor analysis ? *Journal of Econometrics* 132, 169–194.
- Cheung, C. e F. Demers (2007). Evaluating forecasts from factor models for Canadian GDP growth and core inflation.
- Clements, M. e F. Demers (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.
- Dias, F., M. Pinheiro, e A. Rua (2010). Forecasting using targeted diffusion indexes. *Journal of Forecasting* 29(3), 341–352.
- Diebold, F. X. (2015). Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests. *Journal of Business & Economic Statistics* 33(1), 1–24.
- Diebold, F. X. e R. S. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* 13(3), 134–144.
- Eickmeier, S. e C. Ziegler (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting* 27(3), 237–265.
- Ferreira, R. T., H. J. Bierens, e I. Castelar (2005). Forecasting quarterly brazilian GDP growth rate with linear and non linear diffusion index models. *Revista EconomiA* 6(3 dec), 261–292.
- Figueiredo, F. M. R. (2010). Forecasting Brazilian Inflation Using a Large Data Set. *Working Paper Series do Banco Central* (228), 1–57.
- Forni, M., M. Hallin, M. Lippi, e L. Reichlin (2000). The generalized dynamic factor model: Identification and estimation. *The Review of Economics and Statistics* 82(4), 540–554.
- Forni, M. e M. Lippi (2011). The general dynamic factor model: One-sided representation results. *Journal of Econometrics* 163(1), 23–28.
- Gavin, W. T. e K. L. Kliesen (2008). Forecasting inflation and output: Comparing data-rich models with simple rules. *Federal Reserve Bank of St. Louis Review* 90(3), 175–192.
- Gupta, R. e A. Kabundi (2011). A large factor model for forecasting macroeconomic variables in South Africa. *International Journal of Forecasting* 27(4), 1076–1088.
- Hamilton, J. D. (1994). *Time Series Analysis*. New Jersey: Princeton University Press.
- Harvey, D., S. Leybourne, e P. Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting* 13(2), 281–291.
- Harvey, D. I., S. J. Leybourne, e P. Newbold (1998). Tests for Forecast Encompassing. *Journal of Business & Economic Statistics* 16(2), 254–259.
- Hyndman, R. J. e Y. Khandakar (2008). Automatic time series forecasting: The forecast package for R. *Journal Of Statistical Software* 27(3), C3–C3.
- Lima, E. C. R. e P. P. Alves (2011). O desempenho do mercado (Focus) na previsão da inflação: Novos resultados? *Texto para Discussão, Instituto de Pesquisa Econômica Aplicada (IPEA)* jun(1621), 1–25.
- Moser, G., F. Rumler, e J. Scharler (2007). Forecasting Austrian inflation. *Economic Modelling* 24(3), 470–480.

- Newey, W. e K. West (1987). A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3), 703–708.
- R Core Team (2016). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.2, –.
- Schumacher, C. e C. Dreger (2004). Estimating large-scale factor models for economic activity in germany: Do they outperform simpler models? *Jahrbucher fur Nationalokonomie und Statistik* 224(6), 731–750.
- Shmueli, G. (2010, aug). To Explain or to Predict? *Statistical Science* 25(3), 289–310.
- Stock, J. H. e M. W. Watson (1998). Diffusion Indexes. *NBER Working Paper Serie* (6702), 1–55.
- Stock, J. H. e M. W. Watson (2002). Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association* 97(460), 1167–1179.
- Stock, J. H. e M. W. Watson (2005). Implications of Dynamic Factor Models for VAR Analysis. *NBER Working Paper Series* 11467, 1–67.

Apêndice A

Tabela A.1: Descrição das séries macroeconômicas utilizadas

Séries macroeconômicas	Unidade	Identificador	Transf.
1 – Produção			
PIB	R\$	521274780	4
Expedição de caixas - papelão ondulado	Unid.	31873	3
Produção - petróleo - média	Barris/dia	32194	3
Produção física – alimentos	Índice	31845	3
Abate - bovinos - peso das carcaças	Kg (mil)	35921	3
Abate - suínos - peso das carcaças	Kg (mil)	35923	3
Produção - aço bruto	Kg (mil)	36495	3
Produção - cimento	Kg (mil)	37660	3
Produção - caminhões - montados	Unid.	600803224	3
Produção - ônibus - montados	Unid.	600803227	3
Produção - automóveis - montados	Unid.	600803230	3
Produção industrial - bens de capital	Índice	1506214474	3
Produção industrial - bens intermediários	Índice	1506214519	3
Produção industrial - bens de consumo	Índice	1506214564	3
Produção industrial - bens de consumo duráveis	Índice	1506214609	3
Produção industrial - bens de consumo não duráveis	Índice	1506214652	3
Produção industrial - indústria geral	Índice	1506214698	3
Produção industrial - indústria de transformação	Índice	1506214787	3
Utilização da capacidade instalada – indústria	%	33210	0
2 – Preços e indicadores monetários			
M0 - média	R\$	32130	4
M1 - média	R\$	32133	4
IPA origem - prod. agropecuários	Índice	33670	3

Tabela A.1: Descrição das séries macroeconômicas utilizadas

Séries macroeconômicas	Unidade	Identificador	Transf.
IPA origem - prod. industriais	Índice	33672	3
Inflação - INPC	% a.m.	36473	0
Cesta básica - município de São Paulo	R\$	38366	4
Inflação - IPCA	% a.m.	38513	0
Inflação - IGP-DI	% a.m.	39615	0
Inflação - IGP-OG	% a.m.	39617	0
Inflação - INCC	% a.m.	39618	0
Inflação - IPA-EP	% a.m.	39619	0
Inflação - ICV-SP	% a.m.	33239	2
Fatores condicionantes - papel-moeda emitido	R\$	39938	1
Fatores condicionantes - reservas bancárias	R\$	39939	1
Inflação - IPC (FGV)	% a.m.	39621	0
Inflação - IPC (FIPE)	% a.m.	34116	0
3 – Mercado Externo			
Exportações - rentabilidade	Índice	37821	3
Reservas internacionais - liquidez internacional	US\$	32289	2
Serviços e rendas - serviços	US\$	39421	2
Serviços e rendas - rendas	US\$	39424	2
Estados Unidos - índice Dow Jones - fechamento	Índice	39720	3
Estados Unidos - índice NASDAQ - fechamento	Índice	39721	3
Importações - bens de consumo duráveis	Índice	35579	3
Importações - bens de consumo não duráveis	Índice	35580	3
Importações - bens intermediários	Índice	35581	3
Importações - bens de capital	Índice	35582	3
Importações - quantum	Índice	35590	3
Importações - (FOB) - US\$	Índice	36915	3
Exportações - bens de consumo duráveis	Índice	35674	3
Exportações - bens de consumo não duráveis	Índice	35675	3
Exportações - bens intermediários	Índice	35676	3
Exportações - bens de capital	Índice	35677	3
Exportações - quantum	Índice	35690	3
Exportações - (FOB) – US\$	Índice	36916	3
Termos de troca	Índice	1693727000	3
Câmbio contratado - comercial	US\$	32284	0
Câmbio contratado - financeiro	US\$	32287	0
Paridade do poder de compra (PPC) - consumo das famílias	-	548454529	2
Taxa de câmbio - R\$ / US\$ - paralelo - venda - fim período	R\$	35918	3
Relação câmbio efetivo / salário	Índice	39875	3
Conta capital e financeira - conta capital	US\$	39418	0
Conta capital e financeira - conta financeira	US\$	39419	0
Transferências unilaterais correntes	US\$	39531	0
4 – Finanças Públicas			
Execução financeira - receitas	R\$	32226	4
Imposto sobre a circulação de mercadorias (ICMS)	R\$	32110	4
Execução financeira - despesas	R\$	32060	4
Dívida - total - gov. federal e Banco Central - líquida	R\$	32090	4

Tabela A.1: Descrição das séries macroeconômicas utilizadas

Séries macroeconômicas	Unidade	Identificador	Transf.
Imposto inflacionário - real - método 2B	R\$	36719	1
Imposto inflacionário - real - método 2A	R\$	36716	1
Imposto inflacionário - real - método 1B	R\$	36715	1
Imposto inflacionário - real - método 1A	R\$	36714	1
Dívida - total - estados e municípios - líquida	R\$	32088	4
Dívida fiscal - setor público - líquida	R\$	32072	4
NFSP - gov. federal e Banco Central - primário	R\$	1505384299	5
NFSP - estados e municípios - primário	R\$	1505384254	5
Fluxo de caixa da previdência - despesas	R\$	39057	4
Fluxo de caixa da previdência - recebimentos	R\$	1414756693	4
5 – Consumo e vendas			
Vendas de autoveículos*	Unid.	1378	3
Vendas nominais - indústria - SP	Índice	33696	3
Consumo - energia elétrica - comércio	GWh	38357	3
Consumo - energia elétrica - indústria	GWh	38358	3
Consumo - energia elétrica - outros setores	GWh	2027247947	3
Faturamento real - indústria - índice	Índice	33217	3
Inadimplência - índice em t-3	Índice	37852	2
Inadimplência - índice em t-4	Índice	37853	2
SPC - número de registros líquidos	Unid.	40075	2
Consumo aparente - derivados de petróleo - qde./dia	Barril	32035	3
Consumo aparente - bens de capital	Índice	39638	3
Consumo aparente - bens de consumo	Índice	848760088	3
Consumo aparente - bens de consumo duráveis	Índice	848760089	3
Consumo aparente - bens de consumo semi e não duráveis	Índice	848760090	3
Consumo aparente - bens intermediários	Índice	848760086	3
6 – Mercado de trabalho			
Horas trabalhadas - na produção - indústria - SP	Índice	33684	3
Horas pagas - indústria - SP	Índice	33692	3
Pessoal empregado - indústria	Índice	33212	3
Taxa de desemprego - RMSP	%	37655	3
Taxa de desemprego - aberto - RMSP	%	37656	3
Nível de emprego - indústria - SP	Índice	1684652394	3
Salário nominal - indústria - SP	Índice	33694	3
Folha de pagamento - indústria geral	Índice	39625	3
Emprego formal - Índice geral*	Índice	1586	3
Taxa de desemprego - oculto - RMSP	%	37657	3
Horas trabalhadas - indústria	Índice	33208	3
Salário real - médio - indústria - SP	Índice	33690	3
Rendimento médio - real - assalariados - RMSP	Índice	37654	3
7 – Mercado Financeiro			
Índice de ações - Ibovespa - fechamento	% a.m.	31875	0
Taxa de juros - Over / Selic	% a.m.	32241	3
Taxa de câmbio - R\$ / US\$ - comercial - venda - média	R\$	32101	3
Dólar comercial - rendimento nominal	Índice	32220	2
Fundo de ações - rendimento nominal	Índice	32210	2

Tabela A.1: Descrição das séries macroeconômicas utilizadas

Séries macroeconômicas	Unidade	Identificador	Transf.
Índice de confiança do consumidor (ICC)	Índice	40080	3
Índice de expectativas (IEC)	Índice	40082	3
Operações de crédito ao setor público	R\$	1067371625	4
Operações de crédito ao setor privado	R\$	1067371640	4
Operações de crédito ao setor privado - comércio	R\$	1067371633	4
Operações de crédito ao setor privado - imobiliário	R\$	1067371629	4
Operações de crédito ao setor privado - indústria	R\$	1067371627	4
Operações de crédito ao setor privado - outros serviços	R\$	1067371638	4
Operações de crédito ao setor privado - pessoas físicas	R\$	1067371636	4
Operações de crédito ao setor privado - rural	R\$	1067371632	4
Taxa de juros - CDI / Over	% a.m.	32237	3
Taxa de juros - TBF	% a.m.	32244	3
Taxa de juros - TR	% a.m.	32245	2

Fonte: Ipeadata.

Notas: * dados do Banco Central do Brasil, Transformações: 0 sem transformação; 1 deflaciona; 2 primeira-diferença; 3 log-diferença; 4 log-diferença e deflaciona; 5 primeira-diferença e deflaciona.

Comunicação Macropprudencial

Evidências do efeito da comunicação do banco central sobre a estabilidade do sistema financeiro.

*Claudio Oliveira de Moraes¹
Rodolfo Nicolay²*

Abstract

This paper presents a contribution to the analysis of the central bank's communication effects on the capital buffer of commercial banks. The results indicate that facing an indication of increase in interest, banks react with an increase in buffer, suggesting higher risk aversion in relation to the economic environment. The main novelty concerns the communication about restrictive measures of capital. Banks use the buffer itself to meet the new regulatory constraints, so the soundness of the financial system does not increase as expected by regulators.

Keywords: Central bank communication, macroprudential regulation, capital buffer.

JEL Classification: E44, E52, E58

Resumo

O presente trabalho apresenta como contribuição a análise do efeito da comunicação do banco central sobre o *buffer* de capital dos bancos comerciais. Os resultados indicam que frente a uma indicação de elevação de juros, os bancos reagem com um aumento do *buffer*, sugerindo aumento da aversão ao risco em relação ao ambiente econômico. A principal novidade diz respeito ao efeito da comunicação em relação às medidas restritivas de capital. Os bancos utilizam o próprio *buffer* para atender às novas restrições regulatórias, e assim não aumentam a solidez do sistema financeiro, como esperado pelos reguladores.

Palavras-chave: Comunicação do banco central; regulação macropprudencial, *buffer* de capital.

Classificação JEL: E44, E52, E58

¹ Banco Central do Brasil e Universidade Federal Fluminense. Endereço: Sala 434, Campus do Gragoatá – Bloco F – São Domingos, Niterói – RJ – CEP: 24210-350. E-mail: claudiomoraesbcb@gmail.com

² Universidade Candido Mendes, Mestrado em Economia e Gestão Empresarial.

1 – Introdução

Por razões regulatórias os bancos devem manter um determinado nível de capital proporcional aos riscos assumidos, porém, os bancos mantêm um volume de capital superior ao capital regulamentar, o chamado *buffer* de capital. O *buffer* de capital pode ser explicado, do ponto de vista regulatório, como um seguro contra o risco de violar a exigência de capital mínimo, ou seja, os bancos buscam evitar os custos gerados por uma demanda de capital não esperada por parte do regulador. Esta motivação deriva de duas outras razões: (i) os bancos não podem ajustar o capital e o risco instantaneamente; (ii) a violação dos requisitos de capital mínimo regulamentar desencadeia ações de supervisão, que podem levar a uma penalidade com alto custo para os bancos (Stolz e Wedow, 2005; Jokipii e Milne, 2008; Stolz e Wedow, 2011).

Em relação à disciplina de mercado, há uma importante razão para manutenção do *buffer* de capital. O capital adicional visa enviar uma mensagem de solidez para o mercado de tal forma a satisfazer as expectativas de depositante, quando estas não estão garantidas por alguma forma de seguro depósito (Jokipii e Milne, 2008; Nier e Baumman, 2006). Por fim, em relação à estabilidade do sistema financeiro, o *buffer* de capital funciona como amortecedor de capital, pois pode amortecer o efeito de choques adversos sobre os bancos (Nier e Baumman, 2006).

Após a crise *subprime*, o Comitê de Basileia propôs que os bancos centrais adotassem medidas macroprudenciais a fim de garantir a estabilidade financeira. Neste sentido, a análise do comportamento do *buffer* de capital e a comunicação aos bancos comerciais a respeito do *buffer* adequado para manutenção da estabilidade do sistema financeiro surgem como uma nova estratégia de regulação macroprudencial, conforme sugere o Comitê de Basileia para os Bancos Centrais (BCBS, 2010).

A crise desafiou o chamado “*Jackson Hole consensus*”, o qual guiava a prática de diversos bancos centrais que agiam considerando que a estabilidade de preços garantiria a estabilidade financeira (Born, Ehrmann, e Fratzscher, 2012). Uma vez desconsiderada esta hipótese, o banco central passa a ter dois objetivos: estabilidade monetária e estabilidade financeira. Seguindo o teorema de Tinbergen é preciso dois instrumentos independentes para alcançar dois objetivos de política econômica. Desta forma, é preciso tratar a política monetária e a regulação bancária como esferas independentes dentro da atuação do banco central, mesmo que uma esfera possua influência nos objetivos da outra.

O sistema financeiro brasileiro não sofreu os efeitos da crise *subprime*, apesar disso, o Banco Central do Brasil (BCB) foi ativo na condução de comunicados de caráter regulamentar, comunicados na forma de resoluções que visavam afetar o capital regulatório dos bancos. Por outro lado, desde a introdução do sistema de metas para inflação no Brasil o BCB utiliza a ata do COPOM como principal instrumento de comunicação da autoridade monetária, comunicado que visa guiar as expectativas de inflação. Assim, o BCB utiliza duas formas distintas de comunicação para atingir dois diferentes objetivos: estabilidade financeira e estabilidade monetária.

Analisar como a comunicação do BCB afeta o comportamento dos bancos, e em particular o *buffer* de capital, é uma questão relevante para debate atual a respeito das novas estratégias de regulação macroprudencial. Este estudo é uma contribuição a respeito do impacto da comunicação do banco central sobre a estabilidade do sistema financeiro, e assim lança luz sobre um tema pouco explorado pela literatura, a comunicação do banco central enquanto instrumento macroprudencial. O presente

estudo está organizado da seguinte forma: na próxima sessão é apresentada a discussão a respeito dos dois objetivos do banco central; na sessão três são apresentadas as evidências empíricas através dos métodos MQO, GMM e GMM em dois estágios; e na sessão quatro é apresentada a conclusão.

2 – Objetivos do banco central

O banco central possui dois objetivos: estabilidade monetária e estabilidade do sistema financeiro. Desta forma, é possível identificar duas esferas distintas de atuação do BCB no que diz respeito à comunicação enquanto instrumento macroprudencial: (i) a comunicação do banco central enquanto autoridade monetária e seu efeito sobre a estabilidade do sistema financeiro; e (ii) a comunicação do banco central enquanto autoridade reguladora macroprudencial.

2.1 Comunicação do banco central

Segundo Blinder *et al.* (2008), a comunicação da autoridade monetária é fundamental para a gestão da política monetária.³ Em geral, a literatura apresenta evidências sobre o efeito da comunicação da autoridade monetária segundo dois tipos de impactos: o impacto sobre o desempenho econômico e o impacto da comunicação da autoridade monetária sobre os mercados financeiros.⁴

Após a crise *subprime* uma série de trabalhos empíricos surgiram com o objetivo de investigar a relação entre a política monetária e o comportamento dos bancos via canal da tomada de risco⁵. O canal de tomada de riscos se refere à avaliação de como mudanças na política monetária afetam a percepção ou tolerância de risco dos bancos (Borio e Zhu, 2012). De acordo com Altunbas *et al.* (2009) a forma como os bancos assumem riscos sofre influência não só da política monetária, mas também da comunicação do banco central. Com base nesta percepção, Montes e Scarpari (2014) desenvolveram um estudo empírico para o mercado brasileiro no qual identificaram que a comunicação do banco central afeta a tolerância de risco dos bancos via canal das provisões bancárias. De acordo com os autores uma comunicação da autoridade monetária com a sinalização de elevação (ou queda) da taxa de juros induz os bancos a tomar menos riscos (ou mais) de crédito.

Em trabalho a respeito da comunicação do banco central, Born, Ehrmann, e Fratzscher (2012) argumentam que deve-se atentar sobre a diferença entre a comunicação do banco central sobre a política monetária e a comunicação macroprudencial. Porém, os autores não tratam dos efeitos da comunicação do banco central sobre a estabilidade do sistema financeiro, ou seja, a comunicação da autoridade monetária enquanto instrumento macroprudencial. O trabalho se ocupa dos efeitos da comunicação macroprudencial sobre os mercados financeiros, mostrando que tal forma de comunicação é capaz de afetar tais mercados.

2.2 Política Macroprudencial

A principal estratégia para manutenção da estabilidade financeira é o requerimento de capital. Segundo esta regra, os bancos devem manter um determinado

³ Para uma avaliação a respeito da comunicação da autoridade monetária ver: Guthrie e Wright (2000), Ehrmann e Fratzscher (2007), Blinder et al. (2008).

⁴ Para uma avaliação sobre o desempenho econômico ver: Fujiwara (2005), Ullrich,(2008). Para um impacto sobre o mercado financeiro ver: Reeves e Sawicki(2007) e Ehrmann e Fratzscher (2007).

⁵ Para uma análise da literatura empírica do canal de tomada de riscos ver: Altunbas et al. (2009), Gambacorta (2009), e Montes e Peixoto (2014).

volume de capital para fazer frente a todos os tipos de riscos assumidos. Neste sentido, a regra de requerimento de capital está baseada em três hipóteses: (i) hipótese da solvência - o diagnóstico de que o problema da estabilidade financeira é um problema de solvência; (ii) hipótese da mensuração de risco - as estimativas de risco fornecem elementos para uma mensuração do adequado nível de capital. Capital capaz de suportar perdas inesperadas e assim impedir a falência da firma bancária; e (iii) hipótese micropudencial – verificado o cumprimento do nível adequado de capital de cada banco, nos termos da hipótese de mensuração de risco, está garantido a estabilidade do sistema financeiro (Borio, 2011).

Até a crise *subprime* os *policymakers* enfrentaram a gestão da estabilidade financeira como um problema típico de regulação, ou seja, um sistema de regras voltadas para limitar e monitorar o comportamento dos bancos, sem nenhuma associação com a gestão macroeconômica (Woodford, 2010). Neste sentido, a regulação prudencial, principal estratégia regulatória utilizada, tinha seu foco exclusivo na avaliação da qualidade dos balanços e dos empréstimos bancários em um determinado momento, para em seguida determinar se o capital do banco estaria em conformidade com o volume e qualidade dos ativos detidos (Mishkin, 2000).

A crise *subprime* explicitou o fato de que a estratégia regulatória mantida até então não foi suficiente para antever a crise e nem amortecer seus efeitos. Por este motivo, depois da crise, os *policymakers* foram colocados frente à necessidade de mudar a forma de agir e incorporar novos instrumentos macroprudenciais ao conjunto de preocupações da política econômica (Galati e Moessner, 2013). Dentre as alternativas de instrumentos, o *buffer* de capital foi identificado pela literatura como o principal instrumento macroprudencial para gestão da estabilidade do sistema financeiro (Borio, 2011). A proposta é que os bancos aumentem o *buffer* em momentos em que a economia está em expansão para que nos momentos de retração econômica o *buffer* absorva as perdas bancárias. Em suma, o *buffer* quando se comporta de forma anticíclica passa a figurar como o principal instrumento de gestão macroprudencial para estabilidade do sistema financeiro como um todo (BCBS, 2010).

Dois elementos são basilares para a gestão do *buffer* anticíclico, a escolha do indicador que servirá para guiar as decisões dos *policymakers* sobre a demanda de capital adicional (*buffer*) e a comunicação sobre as decisões do *buffer* pelos *policymakers* (BCBS, 2010). Sobre o indicador, o hiato do crédito é considerado pelo Comitê de Basiléia e pela literatura como o principal guia para o *buffer* de capital (Drehmann, Borio, e Tsatsaronis, 2011; Jokivuolle et. al. 2015).⁶ Por outro lado, sobre a comunicação do *buffer*, o Comitê de Basiléia sugere que os bancos centrais levem em consideração o efeito do comunicado do *buffer* sobre as expectativas do mercado, ou seja, sobre o comportamento do próprio *buffer* de capital e em consequência sobre a estabilidade do sistema financeiro (BCBS, 2010). O presente estudo contribui com a literatura de comunicação do banco central ao fornecer evidências sobre o efeito da comunicação da autoridade monetária e da comunicação macroprudencial sobre o *buffer* de capital.

3 – Evidências Empíricas

A análise do efeito da comunicação do banco central sobre o *buffer* de capital é realizada de duas formas. Em primeiro lugar é identificado como a comunicação da

⁶ O hiato do crédito corresponde ao desvio do comportamento da relação crédito/PIB em relação a sua tendência de longo prazo.

autoridade monetária afeta o *buffer* de capital. Em segundo lugar é verificado o efeito da comunicação macroprudencial de caráter normativo sobre o *buffer*. Esta separação tem por objetivo isolar os efeitos dos diferentes tipos de comunicação realizados pelo banco central, neste caso: autoridade monetária e regulador do sistema financeiro.

A comunicação da autoridade monetária é identificada a partir das atas do Comitê de Política Monetária do Brasil (COPOM), que apresenta o comportamento atual da economia e a dinâmica das principais variáveis que influenciam a inflação, além de indicar qual a trajetória esperada para a economia. O BCB iniciou em maio de 2003 a publicação regular, sempre oito dias após a reunião, das atas das reuniões do COPOM. Além disso, as reuniões do COPOM que até 2006 aconteciam uma vez por mês, passaram a ocorrer a cada 45 dias. A base de dados utilizada neste estudo é mensal, logo, quando não há reunião em determinado mês, a avaliação recebida será a mesma que do mês anterior. Esta mesma metodologia foi adotada por Montes e Nicolay (forthcoming) e Montes e Scarparo (2014), considerando que os agentes utilizam sempre a última informação disponível.

A comunicação macroprudencial de caráter normativo é verificada por meio dos comunicados regulamentares de capital, ou seja, por meio da análise das resoluções e circulares publicadas pelo Banco Central, cujo objetivo é alterar a forma de cálculo exigência mínima de capital regulamentar (índice de Basileia). Tais medidas são classificadas como medidas restritivas e medidas liberalizantes. As medidas restritivas são aquelas que restringem a utilização do capital, ou seja, que vinculam determinados ativos ao capital dos bancos. As medidas liberalizantes reduzem a restrição de capital dos bancos, ou seja, reduzem o grau de vinculação dos ativos ao capital dos bancos.⁷

A base de dados conta com 130 observações entre maio de 2003 e fevereiro de 2014. As séries utilizadas na análise empírica são apresentadas a seguir:

BUFFER (*buffer* de capital): Esta série foi construída à partir da série 21424, índice de Basileia. A diferença entre a exigência mínima de capital regulamentar utilizada no Brasil de 11% e o valor da série representa o *buffer*.⁸

ROE (Retorno Operacional sobre o capital): Dados extraídos do Banco Central do Brasil – balancete de instituições financeiras – documento 4010.

O_GAP (hiato do produto): Esta série foi construída a partir da série do PIB acumulado em 12 meses, disponível no site do BCB (série 4190). O hiato do produto é obtido obtendo a diferença entre a série do PIB e sua tendência de longo prazo extraída pelo filtro Hodrick-Prescott.

C_GAP (hiato do crédito): O hiato do crédito é obtido seguindo a metodologia do BCBS (2010). O hiato é a diferença entre a razão crédito/PIB e sua tendência (calculada pelo filtro Hodrick-Prescott). A série da razão crédito/PIB foi retirada do site do BCB (série 17473). Considerando as novas recomendações do BCBS (2010), formalizadas no documento “Guidance for National Authorities Operating the Countercyclical Capital Buffer”, este indicador deve ser utilizado pelos reguladores do Sistema financeiro como critério para a ativação de *buffer* de capital contracíclico. O objetivo é ampliar a capacidade das instituições financeiras de absorver perdas acima do mínimo nos

⁷ O glossário da classificação das atas do COPOM é apresentado na tabela A.3. As medidas de caráter normativas analisadas estão especificadas na tabela A.4.

⁸ Para um detalhamento do cálculo do *buffer* no Brasil ver Tabak et al.(2011).

períodos prósperos do ciclo econômico, para então utilizar a maior reserva de capital nos momentos de estresse.⁹

IR (taxa de juros): Taxa Selic anualizada. A série utilizada é retirada do site do banco central (series 4189).

D_UP (dummy de elevação dos juros), D_DOWN (dummy de redução dos juros) e D_SIG (dummy de sinalização): estas variáveis seguem a metodologia de Rosa e Verga (2007), com a adaptação feita por Costa Filho e Rocha (2009) para as atas do COPOM, para construir o índice de comunicação do banco central acerca da comunicação da trajetória da política monetária no futuro próximo. A variável D_UP recebe valor 1 para as atas que indicam uma elevação dos juros e valor 0 caso contrário; a variável D_DOWN recebe valor 1 para as atas nas quais o BCB indica uma redução na taxa de juros; e D_SIG recebe valor 1 para uma elevação nos juros, -1 para uma redução e 0 quando a política monetária se mostra neutra.

D_LIB (dummy de capital liberalizante), D_RES (dummy de capital restritivo), D_CAP (dummy de sinalização de capital): A variável D_LIB recebe valor 1 para os meses em que há publicação de resolução que indica redução da exigência de capital e valor 0 caso contrário; a variável D_RES recebe valor 1 para os meses em que há publicação de resolução que indica aumento da exigência de capital e valor 0 caso contrário; e D_CAP recebe valor 1 para uma elevação da exigência de capital, -1 para uma redução da exigência de capital e 0 quando não há publicação de nenhuma resolução que trate do capital regulatório.

Como o trabalho utiliza séries temporais em sua análise empírica, uma condição importante para a validade dos resultados é a estacionariedade das séries utilizadas. Para checar se as séries utilizadas possuem um comportamento estacionário em nível foram feitos três testes tradicionais de raiz unitária: Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) e Kwiatkowski-Phillips-Schmidt-Shin (KPSS). Os resultados são apresentados na tabela A.1 no apêndice, e todas as séries foram consideradas estacionárias.

Os resultados empíricos são obtidos a partir de três métodos de estimação: MQO (mínimos quadrados ordinários), GMM (método geral dos momentos) com um estágio e matriz de covariância de Newey-West, e GMM com dois estágios e matriz de covariância de windmeijer. Os resultados do MQO são importantes para a análise de robustez de sinal e significância estatística dos coeficientes. O GMM foi estimado para eliminar os problemas relacionados ao comportamento dos resíduos. Como sabido, séries temporais apresentam problemas tais qual heterocedasticidade e autocorrelação serial, além de endogeneidade das variáveis, e como apontado por Hansen (1982), o GMM é capaz de corrigir estes problemas e apresentar estimações robustas. A escolha dos instrumentos nas estimações por GMM segue a metodologia de Johnston (1984), que recomenda a utilização dos regressores com mais defasagens para garantir a exogeneidade dos instrumentos. Ademais, o teste J é utilizado para verificar a

⁹ De acordo com o BCBS (2010, p. 9): “*The specification of the credit-to-GDP gap has a number of advantages over credit growth. Being expressed as a ratio to GDP, the indicator variable is normalised by the size of the economy. This means it is not influenced by the normal cyclical patterns of credit demand. Being measured as a deviation from its long-term trend, the credit-to-GDP gap allows for the well known secular financial deepening trend. Being a ratio of levels, it is smoother than a variable calculated as differences in levels, such as credit growth, and minimises spurious volatility (no large quarter-to-quarter swings).*”

sobreidentificação nas estimações. As estimações por GMM em dois estágios são utilizadas para obter resultados robustos para pequenas amostras (Windmeijer, 2005).

3.1 Comunicação da Política Monetária

O principal instrumento de política monetária é a taxa de juros de curto prazo. Segundo Montes e Peixoto (2014), a taxa de juros produz efeito sobre o comportamento dos bancos comerciais. De acordo com os autores, diante de um aumento das taxas de juros os bancos reagem com mais provisões, em um claro sinal de aumento de sua aversão ao risco. Assim como as provisões, o *buffer* de capital é uma *proxy* da aversão dos riscos dos bancos. Quanto maior o *buffer*, menor o risco de solvência bancária. Portanto, é esperado que um aumento das taxas de juros (IR) resulte em um aumento do *buffer*. Para analisar o efeito da comunicação do banco central em relação à trajetória da taxa de juros sobre o *buffer* de capital dos bancos comerciais é preciso controlar o efeito da política monetária, por meio da taxa de juros básica (Selic). Outra variável relevante para explicar o *buffer* é o hiato do crédito (C_GAP). Como destaca BCBS (2010), o hiato do crédito é importante para a determinação do *buffer* contracíclico de capital. Em momentos de expansão da economia, os bancos ampliam o crédito e reduzem o *buffer*, e assim aumenta a exposição ao risco de todo o sistema financeiro. Portanto, espera-se que o hiato do crédito tenha o sinal negativo em relação ao *buffer* de capital. O hiato do produto (O_GAP) é uma variável de controle para o ambiente macroeconômico. Espera-se que tenha o sinal negativo, indicando que em um cenário de expansão (contração) econômica os bancos assumem mais (menos) risco e reduzem (aumentam) o *buffer*. Por fim, o retorno sobre o capital (ROE) representa o custo do capital, portanto é esperado que tivesse um sinal negativo.

$$(1) \text{ BUFFER}_t = \alpha_1 + \alpha_2 \text{ ROE}_t + \alpha_3 \text{ O_GAP}_t + \alpha_4 \text{ C_GAP}_t + \alpha_5 \text{ IR}_t + \varepsilon_t$$

$$(2) \text{ BUFFER}_t = \alpha_1 + \alpha_2 \text{ ROE}_t + \alpha_3 \text{ O_GAP}_t + \alpha_4 \text{ C_GAP}_t + \alpha_5 \text{ IR}_t + \alpha_6 \text{ D_UP}_t + \varepsilon_t$$

$$(3) \text{ BUFFER}_t = \alpha_1 + \alpha_2 \text{ ROE}_t + \alpha_3 \text{ O_GAP}_t + \alpha_4 \text{ C_GAP}_t + \alpha_5 \text{ IR}_t + \alpha_7 \text{ D_DOWN}_t + \varepsilon_t$$

$$(4) \text{ BUFFER}_t = \alpha_1 + \alpha_2 \text{ ROE}_t + \alpha_3 \text{ O_GAP}_t + \alpha_4 \text{ C_GAP}_t + \alpha_5 \text{ IR}_t + \alpha_8 \text{ D_SIG}_t + \varepsilon_t$$

As estimações realizadas para as equações de (1) à (4) apresentam os sinais esperados e significância estatística para as variáveis utilizadas para explicar o *buffer* de capital (tabela 1). O teste Ramsey-RESET indica que não há motivos para rejeitar a especificação utilizada no modelo. Ademais, o teste J reportado indica que a escolha dos instrumentos foi adequada e as estimações por GMM estão sobre identificadas.¹⁰

¹⁰ A tabela A.2, no apêndice, apresenta os instrumentos selecionados. É importante destacar que o maior rank utilizado foi 27 (aproximadamente 20% em relação ao número total de observações) e nenhuma variável foi incluída com mais de 6 defasagens para não distorcer os resultados.

Tabela 1: Estimações para o Buffer de Capital (Comunicação da Política Monetária)

	OLS				GMM 1 ESTÁGIO				GMM 2 ESTÁGIOS			
	eq.1	eq. 2	eq. 3	eq.4	eq.1	eq. 2	eq. 3	eq.4	eq.1	eq. 2	eq. 3	eq.4
<i>constant</i>	5.004*** (0.188) [26.647]	4.831*** (0.209) [23.137]	4.951*** (0.177) [27.905]	4.797*** (0.192) [24.990]	5.068*** (0.289) [17.525]	4.823*** (0.341) [14.141]	4.908*** (0.253) [19.422]	4.774*** (0.300) [15.928]	5.386*** (0.519) [10.370]	5.098*** (0.559) [9.124]	5.107*** (0.452) [11.311]	5.110*** (0.527) [9.696]
<i>ROE</i>	-0.025* (0.014) [-1.826]	-0.031** (0.014) [-2.222]	-0.048*** (0.014) [-3.440]	-0.042*** (0.014) [-2.953]	-0.044*** (0.013) [-3.336]	-0.050*** (0.011) [-4.537]	-0.073*** (0.010) [-7.436]	-0.061*** (0.011) [-5.365]	-0.043*** (0.013) [-3.214]	-0.049*** (0.017) [-2.892]	-0.075*** (0.015) [-4.999]	-0.058*** (0.015) [-3.804]
<i>output gap</i>	-6.5E-06*** (7.9E-07) [-8.272]	-7.1E-06*** (8.4E-07) [-8.454]	-7.7E-06*** (7.9E-07) [-9.656]	-7.6E-06*** (8.3E-07) [-9.158]	-7.4E-06*** (7.5E-07) [-9.912]	-7.9E-06*** (7.0E-07) [-11.282]	-8.0E-06*** (5.4E-07) [-14.870]	-8.2E-06*** (6.3E-07) [-12.895]	-7.1E-06*** (1.2E-06) [-5.974]	-7.8E-06*** (1.2E-06) [-6.533]	-7.9E-06*** (8.5E-07) [-9.194]	-7.8E-06*** (1.1E-06) [-7.322]
<i>credit gap</i>	-0.249** (0.106) [-2.347]	-0.271** (0.106) [-2.562]	-0.373*** (0.104) [-3.570]	-0.322*** (0.105) [-3.077]	-0.296** (0.147) [-2.014]	-0.334*** (0.119) [-2.800]	-0.505*** (0.112) [-4.517]	-0.401*** (0.127) [-3.160]	-0.352* (0.192) [-1.827]	-0.418** (0.197) [-2.118]	-0.569*** (0.167) [-3.403]	-0.454** (0.177) [-2.561]
<i>interest rate</i>	0.112*** (0.016) [6.830]	0.123*** (0.017) [7.097]	0.150*** (0.018) [8.322]	0.141*** (0.018) [7.759]	0.124*** (0.020) [6.101]	0.137*** (0.022) [6.177]	0.182*** (0.022) [8.424]	0.159*** (0.024) [6.746]	0.097** (0.042) [2.227]	0.106*** (0.040) [2.625]	0.168*** (0.047) [3.563]	0.128*** (0.046) [2.763]
<i>d_{up}</i>		0.247* (0.135) [1.828]				0.346* (0.184) [1.880]				0.463* (0.246) [1.882]		
<i>d_{down}</i>			-0.548*** (0.134) [-4.09]				-0.731*** (0.170) [-4.299]				-0.741** (0.287) [-2.579]	
<i>d_{sig}</i>				0.254*** (0.078) [3.235]				0.308*** (0.101) [3.043]				0.280* (0.144) [1.943]
aadjusted R ²	0.435	0.445	0.498	0.475	0.435	0.448	0.504	0.485	0.383	0.375	0.465	0.435
F-Statistic	25.822	21.713	26.613	24.315								
Teste F (p-valor)	0.000	0.000	0.000	0.000								
Prob(J-statistic)					0.214	0.339	0.361	0.341	0.372	0.489	0.531	0.507
Rank					16	19	19	19	18	21	21	21
Ramsey-RESET (F-statistic)	0.310	0.259	0.015	0.785								
Ramsey (p-value)	0.579	0.612	0.903	0.377								

Nota: Níveis marginais de significância: (***) denota 0.01, (**) denota 0.05, e (*) denota 0.10. GMM 1 estágio – estimações de GMM um estágio – erro padrão robusto (Newey-West) em parênteses e estatística t em colchetes. GMM 2 estágios – estimações de GMM dois estágios – erro padrão robusto (Windmeijer) em parênteses e estatística t em colchetes. Prob(J-statistic) reporta o respectivo p-valor do teste J

Em relação à comunicação da autoridade monetária, há evidências que esta gera efeitos sobre o *buffer* de capital. Os sinais encontrados indicam, com significância em todos os modelos, que a resposta do *buffer* à comunicação segue a mesma direção da resposta do *buffer* à taxa de juros. Ou seja, quando há indicação de elevação (redução) de juros, os bancos reagem, de forma *forward-looking*, com um aumento (redução) do *buffer*, indicando maior aversão (menor) aversão ao risco em relação ao ambiente econômico prospectivo.

3.2 Comunicação Macroprudencial de Caráter Normativo

Para avaliar o efeito da comunicação de caráter normativo foram utilizadas as mesmas variáveis de controle utilizadas para a análise da comunicação da política monetária: taxas de juros (IR), hiato do crédito (C_GAP), O hiato do produto (O_GAP) e o retorno sobre o capital (ROE). É importante destacar a inclusão de defasagens nas *dummies* de comunicação macroprudencial. Isto é feito para adequar a análise com o fato dos bancos comerciais redefinem suas estratégias a cada trimestre, explicitando as mesmas em seus balanços trimestrais. Por esta razão, é esperado que as medidas de restrições de capital gerem efeitos defasados. A escolha específica das defasagens de três períodos foi feita a partir do método do geral para o específico, considerando a significância estatística dos coeficientes.

Com o objetivo de verificar a relação entre a comunicação de caráter normativo e o *buffer* de capital foram testadas as seguintes especificações:

$$(5) \text{ BUFFER}_t = \beta_1 + \beta_2 \text{ROE}_t + \beta_3 \text{O_GAP}_t + \beta_4 \text{C_GAP}_t + \beta_5 \text{IR}_t + \beta_6 \text{D_LIB}_{t-3} + v_t$$

$$(6) \text{ BUFFER}_t = \beta_1 + \beta_2 \text{ROE}_t + \beta_3 \text{O_GAP}_t + \beta_4 \text{C_GAP}_t + \beta_5 \text{IR}_t + \beta_7 \text{D_RES}_{t-3} + v_t$$

$$(7) \text{ BUFFER}_t = \beta_1 + \beta_2 \text{ROE}_t + \beta_3 \text{O_GAP}_t + \beta_4 \text{C_GAP}_t + \beta_5 \text{IR}_t + \beta_8 \text{D_CAP}_{t-3} + v_t$$

As estimações realizadas para as equações de (5) à (7) apresentam os mesmos sinais e possuem significância estatística para as variáveis de controle (tabela 2). O teste Ramsey-RESET indica que não há problemas de especificação no modelo, e o teste J reportado indica que a escolha dos instrumentos foi adequada e as estimações por GMM estão sobre identificadas.

Os resultados encontrados sugerem que o efeito das medidas regulatórias sobre o *buffer* de capital não geram o resultado esperado. Medidas restritivas geram redução no *buffer* de capital, indicando que os bancos utilizam o próprio *buffer* para atender às novas restrições regulatórias, ou seja, não aumentam a solidez do sistema em termos macroprudenciais. Este resultado é significativo em quase todos os modelos estimados. Por outro lado, medidas liberalizantes geram aumento do *buffer* em todos os modelos estimados, indicando que os bancos não reduzem sua aversão ao risco em função de mudanças regulatórias. Apesar do sinal encontrado nos coeficientes de todos os modelos, só foi identificada significância estatística em um modelo estimado.

Com base nas estimações realizadas, há indícios que a comunicação macroprudencial de caráter regulamentar não atinge os objetivos esperados quando se observa o *buffer* de capital. Ou seja, os bancos utilizam o próprio capital excedente, em relação ao capital regulamentar, para acomodar medidas regulamentares, e desta forma, mantem a decisão de capital interna ao banco. Este resultado corrobora a visão de Nier e Baumann (2006), na qual o *buffer* possui um caráter de acomodação para choques no capital.

Tabela 2: Estimções para o Buffer de Capital (Comunicação Macroprudencial)

	OLS			GMM 1 ESTÁGIO			GMM 2 ESTÁGIOS		
	eq.5	eq.6	eq.7	eq.5	eq.6	eq.7	eq.5	eq.6	eq.7
<i>constant</i>	4.782*** (0.208) [22.941]	4.899*** (0.218) [22.453]	4.835*** (0.207) [23.318]	5.130*** (0.225) [22.841]	5.236*** (0.295) [17.769]	5.357*** (0.268) [19.966]	5.455*** (0.342) [15.948]	5.566*** (0.479) [11.624]	5.452*** (0.472) [11.545]
<i>ROE</i>	-0.024* (0.014) [-1.751]	-0.056 (0.015) [-3.685]	-0.023* (0.013) [-1.746]	-0.035*** (0.006) [-5.466]	-0.034*** (0.009) [-3.579]	-0.028*** (0.008) [-3.587]	-0.038*** (0.011) [-3.415]	-0.030*** (0.011) [-2.684]	-0.037* (0.018) [-1.789]
<i>output gap</i>	-6.4E-06*** (7.8E-07) [-8.171]	-1.8E-05*** (2.6E-06) [-6.745]	-6.2E-06*** (7.9E-07) [-7.856]	-7.1E-06*** (7.0E-07) [-10.100]	-6.2E-06*** (9.6E-07) [-6.499]	-5.6E-06*** (8.6E-07) [-6.529]	-6.1E-06*** (1.1E-06) [-5.423]	-6.2E-06*** (1.3E-06) [-4.777]	-6.2E-06*** (1.7E-06) [-3.735]
<i>credit gap</i>	-0.176 (0.109) [-1.612]	-0.369*** (0.124) [-2.980]	-0.187* (0.108) [-1.725]	-0.472*** (0.116) [-4.072]	-0.229* (0.127) [-1.809]	-0.269*** (0.102) [-2.628]	-0.359** (0.170) [-2.107]	-0.237* (0.143) [-1.664]	-0.345* (0.206) [-1.674]
<i>interest rate</i>	0.130*** (0.018) [7.244]	0.147*** (0.020) [7.519]	0.127*** (0.018) [7.036]	0.105*** (0.020) [5.260]	0.100*** (0.025) [4.007]	0.080*** (0.022) [3.606]	0.076*** (0.029) [2.639]	0.064*** (0.037) [1.752]	0.075* (0.039) [1.942]
<i>d_lib(-3)</i>	0.086 (0.392) [0.220]			0.737*** (0.431) [1.708]			0.401 (0.780) [0.501]		
<i>d_res(-3)</i>		-0.458 (0.289) [-1.585]			-1.243*** (0.615) [-2.020]			-2.012** (0.984) [-2.044]	
<i>d_cap(-3)</i>			-0.315 (0.225) [-1.404]			-0.870* (0.489) [-1.779]			-0.667 (0.874) [-0.762]
<i>adjusted R²</i>	0.456	0.414	0.464	0.343	0.372	0.349	0.333	0.173	0.351
<i>F-statistic</i>	22.090	18.780	22.8253						
<i>Teste F (p-valor)</i>	0.000	0.000	0.000						
<i>Prob(J-statistic)</i>				0.685	0.247	0.384	0.573	0.662	0.413
<i>Rank</i>				27	20	22	22	25	23
<i>Ramsey-RESET</i>	0.163	0.135	0.061						
<i>Ramsey (p-value)</i>	0.687	0.714	0.805						

Nota: Níveis marginais de significância: (***) denota 0.01, (**) denota 0.05, e (*) denota 0.10. GMM 1 estágio – estimções de GMM um estágio – erro padrão robusto (Newey-West) em parênteses e estatística t em colchetes. GMM 2 estágios – estimções de GMM dois estágios – erro padrão robusto (Windmeijer) em parênteses e estatística t em colchetes. Prob(J-statistic) reporta o respectivo p-valor do teste J

4 – Conclusão

A crise financeira internacional impôs à agenda dos *policymakers* a estabilidade do sistema financeiro como elemento fundamental para a estabilidade econômica. A análise dos instrumentos macroprudenciais e a coordenação entre a política monetária e a política macroprudencial ganham relevância para o estudo da macroeconomia pós-crise *subprime*. O presente trabalho visa contribuir com o desenvolvimento de ferramentas macroprudenciais capazes de auxiliar o banco central na tarefa da estabilidade financeira. Neste sentido, foi analisado o efeito da comunicação do banco central, enquanto instrumento macroprudencial, sobre a estabilidade financeira.

A comunicação da política monetária afeta o comportamento do *buffer* de capital. Uma sinalização de aumento (redução) da taxa de juros resulta em um aumento (redução) do *buffer* de capital. Indicando que os bancos aumentam o *buffer* para se proteger de um possível cenário de deterioração econômica ou assumem mais risco em um cenário de bonança econômica. Este resultado sugere que, dado o efeito procíclico da política monetária sobre o *buffer*, os bancos centrais devem considerar a hipótese de coordenação entre a política monetária e a gestão macroprudencial como elemento adicional para a manutenção da estabilidade financeira.

A comunicação macroprudencial de caráter normativo também afeta o *buffer* de capital, porém em um sentido oposto ao esperado. Medidas de caráter restritivo levam a uma redução do *buffer*, sinalizando que os bancos usam o próprio *buffer* para atender às exigências regulamentares e assim não ampliam, de forma macroprudencial, a segurança do sistema financeiro. Este resultado implica na necessidade da revisão da estratégia de capital regulatório. Em suma, uma política macroprudencial focada na estabilidade do sistema financeiro deve ser capaz de atuar de forma direta sobre o *buffer* de capital, como sugere o Comitê de Basileia (BCBS, 2010).

A resposta do *buffer* de capital à comunicação do BCB sugere que a um instrumento de comunicação macroprudencial focada no *buffer* deve ser conduzido de forma semelhante ao utilizado pela política monetária, ou seja, uma comunicação regular e sistemática, capaz de ancorar o comportamento do *buffer* de forma a garantir a estabilidade do sistema financeiro.

Referências

ALTUMBAS, Y., GAMBACORTA, L. e MARQUÉS-IBÁÑEZ, D. (2009), “An empirical assessment of the risk-taking channel”, *Paper presented at the BIS/ECB conference on “Monetary policy and financial stability”*.

BERNANKE, B. (1983), “Non-monetary effects of the financial crisis in the propagation of the great depression.” *American Economic Review*. 73 (3), 257–276.

BASEL COMMITTEE ON BANKING SUPERVISION (2010). Guidance for national authorities operating the countercyclical capital buffer. Basel Committee on Banking Supervision, Bank for International Settlements.

BLINDER, A., EHRMANN, M., FRATZSCHER, M., de HANN, J. e JANSEN, D. (2008), “Central bank communication and monetary policy: a survey of theory and evidence.” *Journal of Economic Literature*, 46(4), 910-45.

BORN, B., EHRMANN M., FRATZSCHER M. (2012), “Communicating About Macroprudential Supervision – A New Challenge for Central Banks.” *International Finance*, 15(2), 179-203.

BORIO, C.; ZHU, H. (2012) “Capital Regulation, Risk-Taking and Monetary Policy: A Missing Link in the Transmission Mechanism?” *Journal of Financial Stability*, 8 (4), 236– 251

BORIO, C. A. (2011) “Rediscovering the Macroeconomic Roots of Financial Stability Policy: Journey, Challenges, and a Way Forward.” *Annual Review of Financial Economics*, 3: 87-117.

COSTA FILHO, A. E. & ROCHA, F. F. (2009). “Comunicação e Política Monetária no Brasil.”. *Revista Brasileira de Economia*, 63, 4.

DREHMANN, M., BORIO, C. e TSATSARONIS, K. (2011), “Anchoring Countercyclical Capital Buffers: The role of Credit Aggregates.” *International Journal of Central Banking*, 7(4), 189-240.

EHRMANN, M., FRATZSCHER, M. (2007), “Communication by central bank committee members: Different strategies, same effectiveness.” *Journal of Money, Credit, and Banking*, 39 (2–3), 509–541.

FUJWARA, I. (2005), “Is the central bank's publication of economic forecasts influential?” *Economics Letters*, 89 (3), 255–261.

GALATI, G., MOESSNER, R. (2013), “Macroprudential Policy – A Literature Review.” *Journal of Economic Surveys*, 27(5), 846–878.

GAMBACORTA, L. 2009. Monetary policy and the risk-taking channel. *BIS Quarterly Review*, December.

GUTHRIE, G., WRIGHT, J. (2000), “Open Mouth Operations.” *Journal of Monetary Economics*, 46(2), 489–516.

HANSEN, L. P. (1982), “Large sample properties of generalized method of moments estimators.” *Econometrica*, 50 (4): 1029–1054.

JOHNSTON, J. (1984), “Econometric Methods, 3.ed.” McGraw-Hill Book

JOKIPII, T. e MILNE, A. (2008), “The cyclical behaviour of european bank capital buffers.” *Journal of Banking and Finance* 32(8), 1440–1451.

JOKIVUOLLE, E., PESOLA, J., VIREN, M. (2015), “Why is credit-to-GDP a good measure for setting countercyclical capital buffer?” *Journal of Financial Stability*, 18 (c), 117–126.

MISHKIN, F.S. (1999),” Global financial instability: framework, events, issues.” *Journal of Economic Perspectives* 13 (4), 3–20.

MISHKIN, F.S. (2000), “Prudential Supervision: why is it important and what are the issue?”, *NBER Working Paper*, 7926

MONTES, G. C., NICOLAY, R. (forthcoming). Central bank communication and inflation expectation. *Journal of Economic Studies*.

MONTES, G. C.; PEIXOTO, G. B. T. (2014),”Risk-taking channel, bank lending channel and the paradox of credibility.” *Economic Modeling*, 39: 82–94.

MONTES, G. C.; SCARPARI (2014), “Does central bank communication affect bank risk-taking?” *Applied Economics Letters*, 22 (9), 37–41.

NIER, E. and BAUMMAN, U. (2006),. “Market discipline, disclosure and moral hazard in banking.” *Journal of Financial Intermediation* ,15,: 332–361.

REEVES, R., SAWICKI, M. (2007), “Do Financial Markets React to Bank of England Communication?” *European Journal of Political Economy*, 23(1), 207–27.

ROSA, C., VERGA, G. (2007), “On the Consistency and Effectiveness of Central Bank Communication: Evidence from the ECB.” *European Journal of Political Economy*, 23(1), 146–75.

STOLZ, S. e WEDOW, M. (2005), “Banks’ regulatory capital buffer and the business cycle: Evidence for german saving and cooperative banks.”: *Banking and Financial Studies*, 249–264.

STOLZ, S. e WEDOW, M (2011), “ Banks’ regulatory capital buffer and the business cycle: Evidence for Germany.” *Journal of Financial Stability*,7(2): 98–110.

TABAK, B.M.,. NORONHA, A.C.B e CAJUEIRO, , D.O.(2011).“Bank Capital Buffers, Lending Growth and Economic Cycle: Empirical Evidence for Brazil.” *Paper prepared for the 2nd BIS CCA Conference on “Monetary policy, financial stability and the business cycle”*.

ULLRICH, K.(2008), “Inflation Expectations of Experts and ECB Communication.” *North American Journal of Economics and Finance*, 19, 93-108

WINDMEIJER, F., (2005). “A finite sample correction for the variance of linear efficient two-step GMM estimators.” *Journal of Econometrics*, 126(1), 25-51.

WOODFORD, M. (2010), “Financial Intermediation and Macroeconomic Analysis.” *Journal of Economic Perspectives*, 24(4), 21–44.

Apêndice

Tabela A.1 – Teste ADF de raiz unitária

ADF					PP				KPSS					
	Leg Lenght	Forma	Teste	Prob	Bandwidth	Forma	Teste	Prob	Forma	Bandwidth	Teste	1%	5%	10%
<i>BUFFER</i>	0	CT	-3.307	0.070	5	CT	-3.415	0.054	CT	8	0.079	0.216	0.146	0.119
<i>ROE</i>	1	CT	-3.405	0.055	5	CT	-3.147	0.100	CT	8	0.075	0.216	0.146	0.119
<i>C_GAP</i>	12	N	-3.909	0.000	2	N	-2.904	0.004	C	9	0.085	0.739	0.463	0.347
<i>O_GAP</i>	1	N	-3.370	0.001	7	N	-2.837	0.005	C	9	0.052	0.739	0.463	0.347
<i>IR</i>	2	CT	-4.578	0.002	8	CT	-3.390	0.057	CT	9	0.150	0.216	0.146	0.119

Tabela A.2 – Instrumentos (estimações GMM)

Política Monetária		
GMM 1	Eq. 1	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -3) O_GDP(-1 TO -3) IR(-1 TO -3)
	Eq. 2	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -3) O_GDP(-1 TO -3) IR(-1 TO -3) D_UP(-1 TO -3)
	Eq. 3	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -3) O_GDP(-1 TO -3) IR(-1 TO -3) D_DOWN(-1 TO -3)
	Eq. 4	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -3) O_GDP(-1 TO -3) IR(-1 TO -3) D_SIG(-1 TO -3)
GMM 2	Eq. 1	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -5) O_GDP(-1 TO -3) IR(-1 TO -3)
	Eq. 2	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -5) O_GDP(-1 TO -3) IR(-1 TO -3) D_UP(-1 TO -3)
	Eq. 3	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -5) O_GDP(-1 TO -3) IR(-1 TO -3) D_DOWN(-1 TO -3)
	Eq. 4	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -5) O_GDP(-1 TO -3) IR(-1 TO -3) D_SIG(-1 TO -3)
Macroprudencial		
GMM 1	Eq. 1	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-3 TO -9) O_GDP(-3 TO -9) IR(-1 TO -3) D_LIB(-4 TO -6)
	Eq. 2	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -4) O_GDP(-1 TO -3) IR(-1 TO -3) D_REST(-4 TO -6)
	Eq. 3	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -3) O_GDP(-1 TO -3) IR(-1 TO -3) D_CAP(-4 TO -9)
GMM 2	Eq. 1	BUFFER(-1 TO -3) ROE(-1 TO -3) C_GAP(-1 TO -3) O_GDP(-1 TO -3) IR(-1 TO -3) D_LIB(-4 TO -9)
	Eq. 2	BUFFER(-1 TO -2) ROE(-1 TO -4) C_GAP(-1 TO -7) O_GDP(-1 TO -4) IR(-1 TO -3) D_REST(-4 TO -7)
	Eq. 3	BUFFER(-1 TO -2) ROE(-1 TO -4) C_GAP(-1 TO -7) O_GDP(-1 TO -4) IR(-1 TO -3) D_CAP(-4 TO -5)

Tabela A.3 – Glossário de classificação das atas do COPOM

Palavras-chave das atas do COPOM	Índice
Banco Central não permitirá que choques de oferta levem a um aumento na taxa de inflação – Manutenção da taxa de juros representa risco não desprezível para o cumprimento da meta (inflação projetada acima da meta) – Riscos para o cumprimento da meta – Impactos inflacionários potenciais de choques de oferta ainda por se materializar – Política monetária deve permanecer vigilante, evitando a propagação de choques e da depreciação cambial – Política monetária firmemente comprometida com a convergência da inflação para as metas – Inflação permanece em patamares elevados/política monetária deve ser firme – Autoridade monetária estará pronta para adotar uma postura ativa, caso se consolide divergência entre inflação projetada e a trajetória das metas – Tendência (da inflação) incompatível com a trajetória de metas – COPOM precisará ser menos tolerante em relação a choques que ameacem tornar a inflação maior do que as metas ou viés de alta	+1
Balanço de oferta e demanda indica ausência de desequilíbrios que possam traduzir-se em	

pressões inflacionárias no horizonte relevante para o regime de metas para a inflação – Quadro geral favorável (apropriado/compatível) à estabilidade de preços (consistente com a meta de inflação), mas fatores de risco precisam ser monitorados (permanecem dúvidas quanto à velocidade de queda da inflação) – Atividade econômica em alta, mas compatível com estabilidade econômica – Comportamento da inflação em linha com as percepções do COPOM – Política monetária se encontra adequada, mas deve ser administrada com cautela (riscos permanecem) – Incertezas recomendam confirmação de quadro favorável – Perspectivas de queda da inflação em direção às metas, mas há riscos de perpetuação da inflação em patamares elevados (trajetória da inflação consistente com as metas) – Aumento da inflação recomenda cautela da política monetária – Probabilidade concreta de a inflação se desviar da trajetória das metas requer cautela adicional da política monetária	0
COPOM decidiu avançar no processo de flexibilização da política monetária – Projeções de inflação abaixo da meta/expectativas compatíveis com as metas/ riscos têm se tornado menos significativos – Consolidação de perspectivas favoráveis para a inflação no médio prazo/COPOM avalia que continuará havendo espaço para quedas adicionais da Selic no futuro – Flexibilização adicional da política monetária deverá ser conduzida de forma parcimoniosa – Cenário benigno para a evolução da inflação (com redução das incertezas/cenário externo favorável) – Atividade econômica em ritmo condizente com as condições de oferta , com baixa probabilidade de pressões sobre a inflação – ou viés de baixa	-1

Tabela A.4 – Classificação das medidas regulatórias de capital.

dez/08	Resoluções 3655 e 3674 e Circular 3425	Liberalizantes
mar/13	Circular 3644	
jul/13	Lei 12838	
jul/14	Circular 3711	
ago/14	Circular 3714	
jan/08	Resolução 3535	Restritivas
jun/10	Circulares 3498, 3568 e 3608	
dez/10	Circular 3515	
nov/11	Circular 3563	
ago/12	Revisão das circulares 3498, 3568 e 3608	
set/13	Resolução 4271	
out/13	Resoluções 4192 a 4195 e 15 Circulares.	

Internet diffusion: killing time prevents crime?

Ilaria Masiero, PhD Candidate at FGV-EESP

Abstract

How does Internet diffusion affect crime rates? I use a panel of state-level yearly data on broadband penetration and criminal activity in the US to shed light on the question. I adopt an instrumental variable approach relying on the fact that the most widespread broadband technologies in the US employ the pre-existing telephone and Cable TV networks to transmit data. The deployment of such networks in early times provides a relevant source of exogenous variation in recent broadband diffusion. Results show that Internet diffusion is associated with a significant and substantial decrease in total and property crime rates, suggesting that a substitution mechanism is at play: online activities crowd out offline activities that more likely lead to crime. A number of tests suggest that my estimates may be interpreted in a causal fashion.

Research area: Applied Microeconomics

Key-words: Internet · Broadband · Crime · Substitution effect

JEL codes: K42, L96

Resumo

A difusão da Internet afeta crime? O presente estudo busca pesquisar o impacto da penetração da Internet sobre as taxas de criminalidade nos EUA. Com o uso de variáveis instrumentais, analisou-se um painel de dados anuais de crime e penetração da banda larga em nível estadual nos EUA. A escolha das variáveis instrumentais se apoia no fato de que as tecnologias de banda larga mais comuns utilizam as pré-existent redes de telefonia e TV a cabo para transmitir dados. Tal uso de redes pré-implantadas fornece uma fonte de variação exógena na difusão recente da banda larga ao longo do território. Os resultados mostram que a penetração da Internet está associada a uma redução significativa e substancial nos índices de criminalidade. Isto sugere que um mecanismo de substituição está em jogo: atividades on-line substituem atividades off-line que mais provavelmente levariam a algum crime. Uma série de testes sugerem que minhas estimativas podem ser interpretadas de uma forma causal.

Área de pesquisa area: Microeconomia Aplicada

Palavras-chave: Internet · Banda Larga · Crime · Efeito de substituição

Classificação JEL: K42, L96

1 Introduction

The Internet has revolutionised many aspects of society and people's lives. Invented at the end of the 1980s, this technology penetrated the US in two waves. The first one involved dial-up connections and started in the early 1990s. The second one began short before the turn of the Millennium, with the introduction of broadband or high-speed Internet technologies. These quickly crowded out dial-up connections, which were down at a 3% penetration rate in 2010, after peaking at about 40% in 2000 (NTIA, 2011). Remarkably, it was only with the introduction of high-speed connections that the Internet reached preponderant levels of penetration in US homes – by 2012, about 70% of American households had broadband Internet at home (NTIA, 2013).

The impact of Internet penetration – especially broadband technologies (see NTIA, 2004; NTIA, 2011) – has been dramatic and far-reaching. In line with this, scholars have put effort in investigating the effect of Internet diffusion on a wide range of outcomes, concerning for instance the labour market (see Kuhn and Skuterud, 2004; Kuhn and Mansour, 2014), marriage rates (see Bellou, 2015), and television viewership (see Liebowitz and Zentner, 2012). One connection that remains relatively unexplored is the possible link between Internet diffusion and crime. Overall, researchers and policy-makers agree that some sort of impact is plausible. However, there is no consensus as to its sign. On the one hand, there is a concern, backed by some results in empirical research (for instance, Bhuller *et al.*, 2011; Anderson *et al.*, 2010), that the abundant supply of sensitive content made accessible by the Internet may positively affect criminal activity, by means of a behavioural or arousal effect. On the other hand, criminological theory suggests that the Internet is capable of affecting crime by altering users' allocation choices. In this sense, the diffusion of the Internet is expected to hinder crime by reason of a substitution effect: online activities crowd out offline activities that more likely lead to crime. The substitution effect may deploy through two closely-connected mechanisms. On the one hand, a catharsis effect (Zimring and Hawkins, 1997): potential criminals substitute actual crime with online activities that inhibit or satisfy their aggressive impulses. On the other hand, a time substitution or voluntary incapacitation mechanism may be at play: having people spend more time online and less time unoccupied on the street *per se* reduces crime. All in all, the substitution effect implies that Internet diffusion negatively affects crime. Evidence in support of this effect has been encountered in previous empirical research (for instance, Dahl and Della Vigna, 2009; Kendall, 2007). Taking a step further, theory has suggested that the substitution effect may be heterogeneous across offense types and/or demographic groups. In particular, it may have a stronger impact on that portion of crime that can be understood as recreational rather than as predetermined and calculated (Zimring and Hawkins, 1997), so that a negative effect of Internet diffusion on this kind of delinquency may be interpreted as the substitution of one diversionary activity with another (for a similar argument see Laqueur and Copus, 2014). The effect may also be more intense on the youth, who typically contribute more than proportionally to crime and are particularly keen on new technologies (see Griffiths and Sutton, 2013).

In summary, theoretical research indicates that Internet diffusion may impact criminal activity both positively, through an arousal effect, and negatively, through the substitution effect. Then, figuring out the overall impact of Internet diffusion on crime is ultimately an empirical matter. Previous empirical literature has investigated some narrowly-defined effects and closely related issues, leaving room for the investigation of the overall effect of Internet diffusion on the most common categories of crime. The main contribution of this research is to shed light on this issue. It is useful to stress that this paper focuses on the most common types of offenses. These do not include cybercrime which, despite being an intriguing phenomenon, is still relatively uncommon. For instance, in 2013 in the US about 6 million larceny-thefts were reported, against 0.26 million cybercrime complaints (see FBI – UCR, 2013; FBI – IC3, 2013).

I exploit the variation in broadband penetration and criminal activity in the US over time and use an instrumental variable (IV) approach with fixed effects to investigate the link between Internet diffusion and crime. The most widespread broadband technologies in the US in the period 2000-2012 technically relied on the pre-existing telephone and Cable TV networks. As a result, the deployment of residential telephone and Cable TV networks in early years may provide a relevant source of exogenous variation in recent broadband diffusion. I investigate this hypothesis through a number of tests, and outcomes provide

support to the validity of this approach. Thus, I instrument the fraction of residential broadband Internet subscribers in each state and year with the rate of residential telephone adoption in 1955 and the share of population employed in the Cable TV sector in 1990, each multiplied by year indicators.

I estimate the model by Two Stage Least Squares (2SLS). Results show a negative and significant effect of broadband Internet on total and property crime rates. All else being equal, a broadband penetration increase by one additional line per 100 people is associated with a 1.4% - 1.5% decrease in crime rates. This effect is rather substantial and should be interpreted as a local average treatment effect (LATE, Imbens and Angrist, 1994). I complement my analysis by presenting the results from estimation by Ordinary Least Squares (OLS) with fixed effects. Again, results uniformly indicate that Internet diffusion is associated with a statistically significant decrease in total and property crime rates, of about 0.6%. 2SLS estimated coefficients are approximately 2.3 times larger in absolute value than OLS estimated coefficients. This discrepancy in the point estimates leads to two considerations. First, the OLS approach is likely to suffer from the omitted variable bias. Secondly, the LATE may be overstating the effect of Internet diffusion for the population at large. Nonetheless, the core result holds that such effect is negative, significant and substantial.

I perform a number of robustness tests to enhance confidence in my IV results and the plausibility of their causal interpretation. To begin with, I run two falsification tests to check that outcomes are not driven by underlying state-specific trends in crime. Then, I estimate my IV model using the limited information maximum likelihood (LIML) estimator, and get to estimates that are remarkably close to the 2SLS ones. Finally, I replicate the analysis using a different source for Internet diffusion data at the state level, as well as at a more disaggregated level. In both cases, results are in line with my baseline estimations.

To develop a deeper understanding of the effect of Internet diffusion on crime, I run two heterogeneity tests along the lines of the hypotheses proposed by previous literature. First, I look at specific types of offenses, and find that broadband Internet diffusion significantly affects each kind of property crime, while it does not seem to have a significant impact on any class of violent offenses. Then, I consider the effect of broadband diffusion on arrest rates for different combinations of age and gender to investigate the hypothesis of heterogeneous effects across demographic groups. None of the estimated coefficients appears to be significantly different from zero in this analysis.

Overall, my results indicate that the effect of Internet diffusion on total and property crime rate is negative and substantial. Relating this outcome to previous theoretical literature, it appears that the substitution effect of the Internet more than outweighs any arousal effect, so that the net impact on crime is negative. This result is aligned with the findings of previous empirical studies analysing the effect on crime of watching violent content over different technological platforms (such as videogames and movies). Outcomes are robust to many falsification and specification tests, enhancing confidence in my baseline results and the plausibility of their causal interpretation. The heterogeneity analyses provide some indication that the substitution effect may have a stronger impact on recreational crime, while heterogeneity of impacts across demographic groups could not be detected.

My research speaks to two policy-relevant debates. Since the beginning of the (broadband) Internet era, policies have been put in place to subsidise its underlying infrastructure and promote its universal deployment, by reason of the positive economic effect predicted for connected communities (see Lehr *et al.* 2006, ITU 2012). However, the impact of Internet diffusion on some social outcomes – such as crime – is ambiguous. Pursuant to this, policy-makers in many countries debate about whether and how internet traffic should be restricted or monitored, to prevent potential adverse effects. The finding that the overall effect of Internet diffusion on crime is negative and significant is relevant to such ongoing debate. In addition, this research relates to the stream of literature investigating the crime drop experienced by many countries in the Developed World since the early 1990s. The drop was unforeseen – quite the opposite, leading experts such as Fox (1996) and DiIulio (1996) had prophesised an explosion in delinquency – and has been object of a great deal of research aimed at pinning down its determinants. There is some consensus on a few factors that have concurred to falling crime rates (see Levitt, 2004). However, given the complexity of the phenomenon, research on this topic is likely never to be exhausted. The present study relates to this literature and suggests that Internet diffusion has played a role in sustaining the crime drop.

The remainder of this paper is organised as follows. Section 2 describes the conceptual framework that motivates the relationship between Internet diffusion and crime. Section 3 introduces the data. In Section 4, I describe the empirical strategy. Section 5 presents the results. I perform robustness and heterogeneity tests in Section 6. Section 7 relates my results to the previous literature and discusses their economic significance. Section 8 concludes.

2 Conceptual framework

2.1 Criminological theory

Routine-changing technologies, such as the Internet, are capable of altering the crime patterns. This is an important conclusion in Routine Activity Theory, a branch in criminology that focuses on how variations in lifestyle (or routine activities) affect the opportunities for – and thus the occurrence of – crime. Routine Activity Theory itself is a sub-field of Crime Opportunity Theory, which maintains that the occurrence of a crime crucially depends on the idiosyncrasies of an immediate situation or context. Crime can be prevented by removing the opportunity to commit it and, if not committed today, it will not necessarily occur tomorrow. Changes in everyday technologies, such as the advent of the Internet, may affect the opportunity for crime and thus criminal activity. The question is: how?

Some criminologists have hypothesised a substitution effect: online activities crowd out alternative offline activities that are more likely to lead to opportunities for crime. The literature seems to identify two relevant and closely-connected mechanisms for the deployment of the substitution effect. On the one hand, a content-driven catharsis mechanism (analogous to the one Zimring and Hawkins, 1997, referred to while studying the effects of mass entertainment on violence): the availability of certain kinds of contents online (for instance: pornography, violence) can inhibit or satisfy aggressive impulses of potential offenders. Then, online activities would serve as a substitute for crime. On the other hand, a time substitution or voluntary incapacitation mechanism may be at play: spending less time unoccupied on the street *per se* reduces the opportunities of committing or being victim of a crime, regardless of the alternative activity that is pursued instead (see Cashmore, 2012; Griffiths and Sutton, 2013). Given that at least until the mid-2000s Internet was mainly accessed from indoor locations, this technology may have negatively affected crime simply by keeping people off the streets for longer. All in all, both mechanisms entail a negative expected impact of Internet diffusion on criminal activity, so that the substitution effect would be negative overall. Taking a step further, some scholars have suggested that this effect may be heterogeneous. Griffiths and Sutton (2013) suggest that the impact may be stronger on the youth, who typically contribute more than proportionally to delinquency and are particularly keen on new technologies.¹ In the words of the authors, “all the time spent online must equate to less time on the street leading to less potential offending time and a smaller population of available victims of violence and robbery”. A further potential source of heterogeneity is that the substitution effect may be more intense on that portion of crime that can be understood as recreational (Zimring and Hawkins, 1997), so that a negative effect of Internet diffusion on this kind of delinquency may be interpreted as the substitution of one diversionary activity with another (see Laqueur and Copus, 2014).

There are other possible ways in which the diffusion of the Internet may affect criminal activity. This technology brought about an unprecedented supply of content, including sensitive material such as violence or pornography, which is available to users with very limited control. Zimring and Hawkins (1997) mention that access to sensitive content may make aggression more likely through a behavioural effect, for instance by altering the emotional state of the audience (the ‘excitement effect’) or by giving them the impulse to re-enact what they have viewed (the ‘ideational’ effect). Similar concerns have been put forward by policy-makers and the public in general, so that the debate as to whether and how internet traffic should be restrained remains open.

In summary, criminological theory has remarked that the diffusion of the Internet is capable of affecting criminal activity both positively, through a behavioural mechanism, and negatively, through the substitution mechanism. Thus, figuring out the actual effect of the Internet on delinquency is ultimately an empirical matter.

¹ According to Wallsten (2013) “people between 15 and 17 spend the most time online, followed by 18-24 year olds”.

2.2 Empirical studies

The literature on the effects of Internet diffusion on delinquency has focused on a specific kind of offenses – sex crimes – reaching non-homogeneous outcomes. Bhuller *et al.* (2011) investigate the impact of Internet use on sex crimes in Norway, in an IV setup. They find that Internet use is associated with a significant increase in sex crimes, likely as a result of increased consumption of pornography. The authors consider among the possible mechanisms both a direct effect of Internet use on the propensity for sex crime (a behavioural effect) and an indirect impact similar to the substitution effect (Internet may affect matching between victim and criminal as its use displaces alternative activities that are more or less associated with sex crimes). Results suggest that, although the indirect effect cannot be ruled out, the net impact of Internet use on sex crime is positive, implying that the behavioural effect is dominant. Kendall (2007) reaches opposite conclusions on the association between Internet use and rape by studying US state-level data over the period 1998-2003. He explains his results in view of a substitution effect through catharsis: criminals are substituting actual rape with pornography. The discrepancy in results between the two studies suggests that underlying differences among geographical contexts may also matter.

Non-Internet-specific empirical research has been carried on in related fields, providing support to the existence of both the behavioural and the substitution effects.

Psychological research provides evidence that subjects exposed to violent material tend to act more aggressively (e.g., Anderson *et al.*, 2010; Anderson, Gentile, and Buckley, 2007). Results are typically derived from experiments in the lab, where a pool of subjects is randomly exposed to some violent content, while the control group watches less violent material. Outcomes usually display statistically significantly heightened cardiovascular activity and hostility measures associated with being exposed to more intense violence. These results are consistent with the existence of a behavioural effect. However, it remains to be established whether higher aggressiveness actually translates into more crimes in the field. Also, because of their experimental setting, these studies cannot provide any indication as to the substitution effect and its relative importance with respect to the behavioural one. Using real-world data, empirical research in economics has analysed the effects on specific kinds of crime of watching sensitive material over platforms other than the Internet. These studies find evidence that both a behavioural and a substitution (through voluntary incapacitation) effects exist. However, the latter more than outweighs the former, so that the net effect on crime is negative. Dahl and Della Vigna (2009) exploit time series variation in the violence level of blockbuster movies from 1995 to 2004 to study the overall short run effect of violent content on violent crime. They estimate that violent movie attendance reduces violent crime in the short term by between 1% and 2%. The authors interpret this finding as a substitution (voluntary incapacitation) effect: potential criminals who choose to go to the movie theatre substitute away from other activities that have higher crime rates. They accept that violent movies may generate a behavioural or arousal effect, evidenced by their finding that there are smaller reductions in violence after more violent movies relative to mildly violent movies. Nonetheless, results show that any arousal effect is dominated by the substitution one. Likewise, Cunningham, Engelstätter and Ward (2011) use a time series approach to estimate the impact of the increased weekly volume of violent video game sale on weekly violent crime. They instrument video game sales with video games ratings in order to only exploit the part of variation in sales due to variation in quality. Their results show that increased violent video games sales produce an overall decrease in violent crime. In line with Dahl and Della Vigna (2009), the authors suggest that this net effect is made up of two components. On the one hand, the finding that violent video games reduce violent crime by a smaller amount than non-violent video games lends support to the existence of a positive behavioural effect. On the other hand, the latter is more than outweighed by a voluntary incapacitation effect, resulting in a negative net impact of violent video games on crime.

In summary, from a theoretical standpoint, the overall effect of Internet diffusion on crime may go in either direction, as the two proposed underlying mechanisms (behavioural and substitution effects) hold opposite signs. Empirical findings provide evidence that both mechanisms exist. Studies analysing the impact of watching sensitive material over platforms other than the Internet mostly agree that the substitution effect dominates the behavioural one. Research on the effects of Internet diffusion itself is scarce, focuses on a very specific kind of crimes (sex crimes) and delivers ambiguous results. Thus, the

overall effect of Internet diffusion on the most common categories of crime remains *a priori* unclear. The purpose of this paper is to empirically address this matter.

3 Data

To investigate the link between Internet diffusion and crime evolution, I exploit the variation in broadband penetration and criminal activity in the US, in a panel data setup. I employ data on crime (the dependent variable), Internet penetration (the explanatory variable of interest), household telephone adoption and employment in the Cable TV sector (the instruments), as well as on a number of demographic and socio-economic controls. Data is organised in a panel covering a thirteen-year period (2000-2012) for 51 territories (50 states and the federal district). The information used in this study comes from several sources.

Crime data

State-level, annual data on crime occurrences are taken from the Uniform Crime Reports (UCR) issued by the Federal Bureau of Investigation (FBI). These data have been used in mainstream crime literature, for instance by Kelly (2000) and Donohue and Levitt (2001). Data refer to seven felony offenses entailing both violent and property crimes. Violent crimes consist of murder and non-negligent manslaughter, forcible rape, robbery and aggravated assault; property crimes include burglary, larceny-theft and motor vehicle theft.² The FBI selects these crime categories because they are the most likely to be reported and to occur with sufficient frequency to provide an adequate basis for comparison.³ Total, violent and property crime rates per 100,000 residents represent the main dependent variables in the analysis. I use the logarithmic transformation because I expect the variation in crime to be proportional to its initial level.

In a robustness check, I replicate the analysis at a more disaggregated level. Local data on crime are from FBI's UCR too.

Finally, in a heterogeneity test, I use as dependent variables arrest rates for different combinations of age and gender. Data on arrests by age and sex are from the FBI's UCR.

Internet data

The explanatory variable of interest must reflect Internet penetration. To the best of my knowledge, there are no available yearly, state-level data on the diffusion of Internet technologies other than broadband. Thus, the data I use actually picture the diffusion of broadband Internet. Although this is not ideal, it shall be kept in mind that, starting from their introduction short before year 2000, broadband platforms have largely replaced other connections (NTIA, 2011; NTIA, 2013). In addition, the impact of broadband technologies has arguably been more intense than that of previous technologies (see NTIA, 2004; NTIA, 2011), and it was only with the introduction of broadband platforms that the Internet reached preponderant levels of penetration in US homes. Thus, it does not seem unsound to approximate the effect of the Internet diffusion in the period 2000-2012 using broadband penetration. Still, I perform a robustness check using a different Internet data source to get an indication as to whether this approximation is reasonable (see Section 6.1.3).

In my baseline analysis, I use data on broadband diffusion from the Federal Communications Commission's (FCC) Statistical Reports on Broadband Deployment. These data have already been employed in the literature. For instance, Bellou (2015) uses them to examine the impact of broadband Internet penetration on marriage rates in the US. The FCC's Statistical Reports on Broadband Deployment provide information on subscribership to Internet access services gathered through the Form 477, which qualifying providers are required to file twice a year. Information includes data on the number of broadband or high-speed lines (connections to end-user locations that deliver services at speeds exceeding 200 kilobits per second in at least one direction) by type of user (residential or business) and per state. I focus on the residential segment because the literature has suggested that Internet availability may have an impact on crime by affecting people's spare time allocation choices. Data on residential

² For felonies' definitions, see this [link](#).

³ FBI's [website](#).

broadband penetration has been recorded half-yearly since December 2000.⁴ Because of this data limitation, my analysis will cover the period starting from 2000 although broadband deployment actually started at the end of the 1990s. I use for each year the second semester's report, portraying the situation as of the end of the year. Data are not available for Hawaii in the period 2000-2005 and for Wyoming in 2000. My explanatory variable of interest is built as the number of residential broadband lines per 100 people in a certain state and year.

In a robustness check at both state and local level, I employ as an alternative source for Internet diffusion data the Internet and Computer Use Supplement data collected by the US Census as part of its Current Population Survey (CPS). This source has been used in previous studies. For instance, Kuhn and Mansour (2011) used it to investigate the impact on unemployment duration of using the Internet in the job search strategy. CPS Internet data were only collected sporadically during the period of interest, which is why this is not my baseline source. I focus on a variable indicating whether the household has an Internet connection, regardless of the connection technology – the survey question is: “Does anyone in this household connect to the Internet from home? Yes/No”. Over the period of interest, this question was asked in years 2000, 2001, 2003, 2007, 2009 and 2010. The fact that this variable is not broadband-specific allows me to get an indication as to whether the effects of broadband penetration can soundly approximate those of Internet diffusion at large.

Further data

I use data on state-level residential telephone adoption and employment in the Cable TV sector in early years to build my instrumental variables. Data on residential telephone adoption in 1955 are from the 1956 City and County Data Book collected by the US Bureau of Census. Information is unavailable for Hawaii and Alaska. In a robustness test, I replicate the analysis at a more disaggregated level. Local data on residential telephone adoption in 1986 are from the CPS. State- and local-level data on employment in the Cable TV sector in 1990 are available by the US Bureau of Labor Statistics. I use US Standard Industrial Classification code 4841, referring to “Cable and other pay TV services”.

I include a number of controls in my regressions. Data about states' characteristics (area, population, age and race structure) are from the US Bureau of the Census. State-wide annual averages of unemployment rates come from the US Bureau of Labor Statistics. Data on per capita personal income are available by the US Bureau of Economic Analysis. I consider the logarithmic transformation of this variable throughout the analysis.

Descriptive statistics

Table 1 displays the mean and standard deviation in years 2000 and 2012 for the state-level, time-varying variables described above, as well as their percentage variation in mean over the said period. The pace of broadband Internet diffusion reflected by the FCC data has been exorbitant over the period 2000-2012, at more than 4000%. On the other hand, the increase in Internet penetration (all technologies) portrayed by the CPS data has been less outstanding as dial-up connections were already quite diffused in 2000, when they started being replaced by broadband platforms. All crime rates have substantially fallen since the beginning of the Millennium. In fact, much of the Developed World has experienced a sharp drop in crime rates since the early 1990s.

4 Empirical strategy

4.1 Baseline model

I estimate the effect of broadband Internet penetration on crime rates evolution in the US using a panel that covers a thirteen-year period (2000-2012) for 51 territories (50 states and the federal district). I employ an IV approach, which is summarised by the following equations:

- (1) $y_{st} = \beta_0 + \beta_1 bbrate_{st} + \beta_2 X_{st} + \rho_t + \theta_s + \varepsilon_{st}$
- (2) $bbrate_{st} = \beta_0 + \beta_1(\rho_t * phonerate_s) + \beta_1(\rho_t * cableTV_s) + \beta_2 X_{st} + \rho_t + \theta_s + \varepsilon_{st}$

⁴ Until the second release of 2004, Residential and Small Business are recorded as a single user type. Afterwards, the Residential and Business segments are recorded separately.

Where the subscripts s and t denote state and year respectively; y_{st} is the crime rate per 100,000 residents; $bbrate_{st}$ is the number of residential high-speed Internet lines per 100 residents; $phonerate_s$ is the number of residential phone lines per 100 residents in 1955; $cableTV_s$ is the number of Cable TV employees per 100,000 residents in 1990; ρ_t and θ_s are year and state fixed effects respectively; and X_{st} is a vector of demographic and socio-economic controls, namely: population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. In my baseline analysis, I estimate the model above by 2SLS, whereby Equation (1) represents the second stage and Equation (2) represents the first stage. I use the robust matrix variance estimator with standard errors clustered at the state level throughout the analysis.

4.2 IV approach

Broadband diffusion is unlikely to be random with respect to the determinants of crime, even conditional on the model covariates. For example, states where cities are growing faster may envisage faster broadband penetration and higher crime growth rates. As a consequence, standard OLS estimation of the relationship between Internet diffusion and crime would likely suffer from the omitted variable bias, and estimates would reflect a measure of association rather than a causal effect. To (more) plausibly attach a causal interpretation to estimation results, it is necessary to somehow filter the overall variation in broadband penetration and only use the part of it that is exogenous with respect to the determinants of crime evolution. The IV approach aims at implementing this intuitive strategy.

Instrument definition

I use as instrumental variables for my analysis the number of residential telephone lines per 100 residents in 1955 and the number of people employed in the Cable TV sector per 100,000 residents in 1990. I interact the instruments with the time fixed effects because they are time invariant, while the endogenous variable (crime rate) is not. More specifically, for each state and year, I instrument $bbrate_{st}$, the fraction of population with a residential broadband Internet subscription, with $(\rho_t * phonerate_s)$ and $(\rho_t * cableTV_s)$, respectively the fraction of population that had a residential telephone in 1955 and the fraction of population that was employed in the Cable TV segment in 1990, multiplied by year indicators. In the next paragraphs I motivate the validity of my instruments in terms of both relevance and exogeneity.

Instrument motivation

Broadband connection can be provided over different platforms, which entail different technologies. As shown in Figure 1, at least until the end of the years 2000s, the two most widely diffused platforms for residential clients in the US were the Digital Subscriber Line (DSL) and the Cable Modem – whereas the Mobile Wireless technology took over starting from year 2011. The DSL and Cable Modem platforms rely on the physical networks of previously-existing communication technologies: the telephone's copper lines and the Cable TV's coaxial cables. This fact suggests that residential broadband Internet diffusion has followed along the penetration path of these earlier communication technologies. Also, this line of reasoning relates to a more general notion in previous literature (see Stevenson, 2006; Bellou, 2015) suggesting that the adoption of communication technologies by households follows some long-standing patterns driven by state-specific factors that are rather stable over time. My IV definition starts from the hypothesis that the deployment of residential telephone and Cable TV networks in early years represents a good predictor for recent broadband diffusion.

To support my reasoning, I discuss some technological features of the DSL and Cable Modem technologies. As explained by the FCC, “DSL is a wireline transmission technology that transmits data faster over traditional copper telephone lines already installed to homes”.⁵ “With the advent of the modem, telephone networks were the first networks to provide Internet access. After all, millions of homes were already ‘wired’ with twisted-pair copper lines”.⁶ Initially, dial-up Internet used the same analogue network designed for voice to deliver Internet access at lower speeds. To offer high speed access, the network needed to be upgraded. Then, the DSL technology envisaged a cost advantage

⁵ FCC's [website](#).

⁶ FCC (2010).

because, since it is “deployed over the same existing twisted-pair copper network used to deliver telephone service, it benefits from sunk costs incurred when first deploying the telephone network”.⁷ Similar considerations hold for the Cable Modem technology. According to the FCC, this platform “enables cable operators to provide broadband using the same coaxial cables that deliver pictures and sound to your TV set”.⁸ “Cable systems were originally constructed to provide one-way video signals [...]. In the 1990s with the advent of the Internet and passage of the 1996 Telecommunications Act, cable companies began upgrading their networks to provide the two-way transmission capabilities required for Internet data traffic [...]. Such upgrades were seen as attractive since millions of homes were already ‘wired’ with high capacity coaxial cable”.⁹

Data seem to confirm that telephone and Cable TV early deployment is relevant to the pattern of recent broadband diffusion. Figure 2 displays actual state-level broadband Internet penetration rates plotted against those predicted by the rate of residential telephone ownership in 1955 and the fraction of population employed in the Cable TV sector in 1990, for years 2000, 2004, 2008 and 2012. In each year, the state dots are clustered around the 45 degree line (at different penetration levels). Although only four cases are displayed as an example, an analogous outline extends to all sample years. The predictive power of the historical telephone and Cable TV diffusion patterns for broadband Internet penetration is remarkable and confirmed by the first-stage estimation results (see Table 2). This outcome was expected, given that broadband technologies rely on the pre-existing networks of the telephone and Cable TV. What is more, these results are in line with the more general notion that some unobserved, state-specific factors drive the penetration pattern of all residential communication technologies and are rather stable over time (see Stevenson, 2006; Bellou, 2015). Then it seems reasonable that the deployment of residential telephone and Cable TV networks in early years is relevant to recent broadband diffusion.

Further than relevant, a good instrument in this context must also be exogenous with respect to the time-varying determinants of recent crime evolution. In general, instrument exogeneity cannot be tested. At best, having more instruments than endogenous variables, the Hansen J test for overidentifying restrictions can be performed. This test assumes that at least one of the instruments is exogenous and verifies whether all instruments are exogeneous. Results from the Hansen J-test are displayed in the last two rows of Table 3. They lend some support to the validity of the exogeneity condition for my instruments. To further enhance the plausibility that this condition is satisfied, the following considerations hold: (i) I take the value of my instrumental variables in early times (1955 for the telephone and 1990 for Cable TV) in order to help rule out any reverse effect of the dependent variable (crime rates in the period 2000-2012) on the instruments – as well as to have a higher level of heterogeneity among states; (ii) both telephone and Cable TV technologies had virtually reached their full deployment by the turn of the Millennium, when the broadband Internet diffusion began. This helps rule out possible direct effects of the instruments on recent crime rates; (iii) among the factors that may be correlated with both broadband and crime evolution, it is important to distinguish between time-varying and time-invariant ones. As argued above, I do not rule out that the pattern of broadband Internet adoption reflects some state-specific features, such as housing or cultural characteristics, which may well be correlated with recent crime patterns. However, given that these characteristics are stable over time, their effect will be absorbed by the state indicators. As regards time-varying factors, I perform falsification tests to check that my results are not driven by underlying state-specific trends (see Section 6.1.1).

In summary, the crucial hypothesis in my IV approach is that the deployment of the telephone and Cable TV networks in early years provides a relevant source of exogenous variation in broadband penetration. Although this hypothesis cannot be verified with certainty, outcomes from a number of tests reinforce my confidence in its validity.

⁷ FCC (2010).

⁸ FCC’s [website](#).

⁹ FCC (2010).

5 Results

In my baseline analysis I estimate the model by 2SLS, whereby Equation (1) represents the second stage and Equation (2) represents the first stage. I begin by presenting the first stage outcomes. Second stage results follow suit. I also report the results from OLS estimation of Equation (1) and display summary outcomes from the reduced form regression of the dependent variables directly on the instruments.

5.1 First stage estimates

First stage estimation results reported in Table 2 suggest that my instruments are relevant. Telephone and Cable TV deployment variables are highly correlated with crime rates after controlling for the exogenous covariates. As expected, the estimated coefficients are positive: a higher level of penetration of residential telephone and Cable TV technologies in early years is associated with greater broadband Internet diffusion in recent times. The coefficients are highly significant starting from year 2003. Plausibly, in the initial period broadband Internet was not widespread enough for a significant effect to be detected. Starting from 2003, the effect is precisely captured, also thanks to the utter prevalence of DSL and Cable Modem technologies against other broadband platforms (see Figure 1). Around the end of the sample period, again significance is eroded, especially as far as the telephone adoption instrument is concerned. This may be due to the relatively higher expansion of Mobile Wireless broadband connections – which do not rely on the telephone’s network – with respect to the DSL and Cable Modem technologies. The F-statistic for a test on whether the excluded instruments are significant is close to 10, reinforcing my confidence that the instruments are not weak.

5.2 IV and OLS estimates

Turning on to second stage estimation, outcomes are displayed in the first three columns of Table 3. Results point to a negative and significant effect of broadband Internet on total and property crime rates. All else being equal, a broadband penetration increase by one additional line per 100 people is associated with a 1.4% - 1.5% decrease in total and property crime rates. On the other hand, the estimated effect on violent crime is not significant. These coefficients should be interpreted as a LATE: by construction of my IV model, they represent the effect of broadband Internet deployment on crime rates for that sub-population (‘compliers’) for which broadband penetrated according to the patterns of pre-existing telephone and Cable TV networks. The last two rows in columns 1 to 3 of Table 3 report the Hansen J statistic and associated p-value. As mentioned, the Hansen J test assumes that at least one of the instruments is exogenous and tests whether all instruments are exogenous. The joint null hypothesis is that the over-identifying restrictions are uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. The test does not reject the null hypothesis, thus it does not rule out that my instruments are indeed exogenous.

The last three columns in Table 3 display the outcomes from OLS estimation of Equation (1). Again, results envisage a negative association between broadband diffusion and total and property crime rates. The estimated coefficients indicate that, all else being equal, a broadband penetration increase by one additional line per 100 people is associated with a 0.6% decrease in crime rates. OLS estimated coefficients are about 2.3 times smaller in absolute value than 2SLS estimated coefficients. Two considerations help explain this discrepancy. First of all, the OLS approach is likely to suffer from the omitted variable bias because broadband diffusion is probably non-random with respect to the determinants of crime (conditional on other covariates). In particular, it seems to be the case that OLS estimates do not account for some heterogeneity that sets back the real effect of Internet penetration on crime rates. A consistent explanation for this bias would be, for instance, that states with faster growing cities record both faster Internet penetration rate and slower crime declines. The second consideration is that some compliers’ characteristics relevant to crime evolution may not be homogenous with respect to the overall population, so that the LATE may be overstating the effect of Internet diffusion on crime for the population at large. OLS estimates do not suffer from this kind of bias and thus may more closely reflect the effect of Internet diffusion on crime for the population at large. All in all, considering the results from the two approaches jointly suggests that the effect of Internet diffusion on total and property crime is negative and significant.

Relating outcomes to previous theoretical literature, a negative net impact of broadband Internet diffusion on crime suggests that any arousal effect is dominated by the substitution effect through which online activities crowd out alternative offline activities that are more likely to lead to crime. In connection with to previous empirical research, my finding that the net effect of broadband diffusion on crime is negative is in line with the conclusions reached by those studies analysing the impact on crime of technological platforms other than the Internet. In addition, my results are not inconsistent with the notion of heterogeneous effects by type of offense, as broadband Internet diffusion seems to affect property but not violent crime. I further assess this hypothesis in Section 6.2.1.

5.3 Reduced form relation

Figure 3 reports summary results from the reduced form estimation, which entails regressing crime rates directly against the instruments and further covariates.¹⁰ The reduced form estimates envisage a negative and significant relation between telephone adoption rate and total and property crime rates. Importantly, the reduced form relation parallels the pattern displayed by the first stage regression, whereby the link between the variables is not precisely estimated in the first years of the sample, it becomes significant in central years, and finally fades away towards the end of the sample period. This parallel outline reinforces the plausibility that this significant reduced form association in fact reflects the underlying relationship between broadband Internet penetration and crime, and is motivated by the link between the early deployment of the telephone network and recent broadband Internet diffusion.

On the other hand, the reduced form estimation detects no significant relationship between crime rates and the share of population employed in the Cable TV sector in 1990, as displayed in Figure 4.

6 Robustness and heterogeneity tests

6.1 Robustness tests

In this section, I perform additional tests to enhance confidence in my IV results and the plausibility of their causal interpretation.

To begin with, I run two falsification tests to further assess the validity of my instrumental approach. Secondly, I estimate the IV model using the limited information maximum likelihood (LIML) estimator. Then, I replicate the baseline analysis at both state and local level using a different source for Internet diffusion data.

6.1.1 Falsification tests

An important requirement for my IV setup to be valid is that the instruments are unrelated to the time-varying determinants of crime evolution. In fact, if the exogeneity condition were violated, no causal interpretation could plausibly be attached to the results, as they would just reflect the relationship between Internet diffusion and the underlying state-specific trends in crime. I assess this concern by performing two falsification tests. In the first one, I exploit the timing of the advent of broadband Internet to increase confidence that the instruments are unrelated to underlying trends in crime across states. In the second test, I allow crime evolution to be related to state-specific trends in the exogenous covariates, to check whether the significant effect detected in my baseline analysis holds or is absorbed by such trends.

Broadband penetration occurred starting from the late 1990s – even though the analysis takes 2000 as the baseline year due to lack of available data for the previous period. If the instruments are exogenous, there should not be a significant association between them and crime rates prior to the late 1990s. On the other hand, a systematic relationship would suggest that the exogeneity condition is violated. I take 1996 to be the year in which broadband penetration started because literature (for instance, Bellou, 2015) indicates the 1996 Telecommunications Act – deregulating the broadcasting market and promoting competition in the telecommunications industry – as the turning point in broadband history. This falsification test studies the association between the instruments and crime rates over the period 1990-2012. Summary results for the telephone adoption instrument are presented in

¹⁰ Extensive results are available upon request.

Figure 5.¹¹ Remarkably, the estimated coefficients for total and property crime associated with telephone adoption start being consistently significant in 1996 and highly so (at the 1% level) in 1997. In addition, the share of people employed in the Cable TV sector in 1990 is not significantly associated to crime rates in the pre-broadband years. These results enhance confidence that my instruments are exogenous with respect to the underlying state-specific trends in crime.

A limitation to this test is that, before broadband technologies, dial-up Internet diffused to US households starting from the early 1990s – its penetration rate was at about 40% in 2000 (NTIA, 2011). Like the DSL platform, the dial-up technology physically relies on the telephone network. Thus, a finding of some negative and significant effect in the pre-broadband would not totally rule out the exogeneity of my instruments, as it may reflect the dial-up Internet effect on crime. In particular, this may explain why the estimated coefficients for total and property crime in year 1994 are significant. Still, starting from the beginning of the broadband era (1996-1997), the effect is consistent and more precisely estimated, in line with the facts that (i) the impact of broadband technologies has arguably been more intense than that of dial-up connections (see NTIA, 2004; NTIA, 2011), and (ii) it was only with the introduction of broadband technologies that the Internet reached preponderant levels of penetration in US homes. It is useful to stress that the said limitation only applies to this specific test and does not extend to the overall analysis. In fact, my purpose is to investigate the impact of the diffusion of Internet (regardless of the technology) on crime. Therefore, as long as my instruments are capturing the effect of the diffusion on crime of some kind of Internet technology, my analysis is not dampened. As explained in Section 3, I am limited by data availability issues to approximate overall Internet diffusion by broadband Internet diffusion. This does not seem unreasonable given the above-mentioned points (i) and (ii) together with the fact that broadband technologies have quickly crowded out dial-up connections since their introduction. Still, I provide some indication as to the reliability of this approximation by replicating the analysis using a different source of Internet diffusion data (the CPS) that covers all types of Internet connections for a more limited number of years (see Section 6.1.3).

All in all, this first falsification test reinforces my confidence that the instruments are exogenous.

I further assess this condition with another test. Following Duflo (2001), I include as further covariates in Equation (1) the interactions between baseline year (2000) exogenous covariates and either a linear time trend or time dummies. This falsification test allows crime evolution to be related to state-specific trends in the exogenous covariates. The purpose of this analysis is to check whether the significant effect detected in my baseline analysis holds even in this setting.

Equation (1) becomes:

$$(3) \quad y_{st} = \beta_0 + \beta_1 bbrate_{st} + \beta_2 X_{st} + t \sum_j \lambda_j x_{(s,2000),j} + \theta_s + \varepsilon_{st}, \text{ with a time trend; or}$$

$$(4) \quad y_{st} = \beta_0 + \beta_1 bbrate_{st} + \beta_2 X_{st} + \rho_t \sum_j \lambda_j x_{(s,2000),j} + \theta_s + \varepsilon_{st}, \text{ with time dummies.}$$

Where $x_j, j = (1, \dots, J)$, are the covariates I am using in my baseline model.

Table 4 displays outcomes from the 2SLS estimation of Equation (3) – columns 1 to 3 – and Equation (4) – columns 4 to 6. Remarkably, these outcomes show that the negative and significant relationship between broadband diffusion and total and property crime rates is not accounted for by underlying state-specific trends in the determinants of crime. In fact, the estimated coefficients for total and property crime remain negative and significant. Point estimates from the estimation of Equation (3) are very close to my baseline results, while those from Equation (4) are a bit larger in absolute value. These results also suggest that the effect of Internet diffusion may be significant also on violent crime rates, once underlying state-specific trends in the determinants of crime are accounted for. Overall, the analysis reinforces my confidence that the instruments are exogenous.

6.1.2 LIML estimation of the IV model

The 2SLS estimator is consistent but biased, and the bias is an increasing function of the number of instruments. In my IV model, I have one endogenous variable and twenty-four instruments. Therefore, biasedness of estimates may be an issue. To address this concern, I follow Angrist and Pischke (2009) and estimate my IV model through LIML. The LIML estimator is less precise than the 2SLS one, but it is

¹¹ Extensive results are available upon request.

approximately median-unbiased for over-identified fixed effects models. Therefore, similar outcomes from the 2SLS and LIML estimations would suggest that biasedness in 2SLS estimates is not much of a concern. This appears to be the case, as shown in Table 4. LIML results confirm a negative and significant relationship between broadband diffusion and total (-1.6%) and property (-1.7%) crime. In terms of sign and significance, these outcomes are identical to results from the 2SLS estimation. In terms of exact point estimates, LIML estimated coefficient are slightly higher in absolute value with respect to 2SLS ones, but still very close. Overall, these outcomes suggest that biasedness of my baseline IV estimates does not raise serious concerns.

6.1.3 Analysis using CPS Internet data

I further test my results by replicating the analysis with a different source of data for Internet diffusion. I use the Internet and Computer Use Supplement data collected by the US Census as part of the CPS. This source has been widely used in previous literature on the effects of Internet penetration on socio-economic outcomes (for instance, Kendall, 2007; Kuhn and Mansour, 2011). I focus on a variable indicating whether the household has an Internet connection, regardless of the connection technology – the survey question is: “Does anyone in this household connect to the Internet from home? Yes/No”. This variable is available in six years over the period 2000-2012 (namely: 2000, 2001, 2003, 2007, 2009 and 2010), and has been used in previous related literature (for instance, Kendall, 2007).

Equation (1) and Equation (2) become, respectively:

$$(5) \quad y_{st} = \beta_0 + \beta_1 HHint_rate_{st} + \beta_2 X_{st} + \rho_t + \theta_s + \varepsilon_{st}$$

$$(6) \quad HHint_rate_{st} = \beta_0 + \beta_1(\rho_t * phonerate_s) + \beta_1(\rho_t * cableTV_s) + \beta_2 X_{st} + \rho_t + \theta_s + \varepsilon_{st}$$

Where $HHint_rate_{st}$ is the natural logarithm of the percentage of households that have an Internet connection in state s and year t .

The purpose of this specification check is twofold. First, since CPS Supplement has been widely used by previous literature, it is sensible to replicate my analysis using this information to check whether my baseline results are confirmed. Secondly, this analysis provides some indication as to whether it is reasonable to approximate the effect of the Internet diffusion at large using broadband penetration. In my baseline analysis, I have assumed that such approximation is sound based on the facts that (i) starting from their introduction around year 2000, broadband platforms have largely replaced dial-up connections (NTIA, 2011; NTIA, 2013); (ii) the impact of broadband technologies has arguably been more intense than that of dial-up connections (see NTIA, 2004; NTIA, 2011), and (iii) it was only with the introduction of broadband technologies that the Internet reached preponderant levels of penetration in US homes. In line with this, the Internet diffusion variable I employ in my baseline analysis only refers to broadband technologies. This choice was also dictated by data limitation issues, as overall Internet diffusion data are only available by the CPS on sporadic years. In this specification check, however, I use CPS data for the available years and consider a variable that covers all types of Internet connections. The idea is that if the analysis were to generate substantially different results with respect to my baseline outcomes, this could imply that the effect on crime of broadband technologies is considerably different with respect to the effect of the Internet at large.

The first three columns in Table 5 present the second stage results of the IV model. For completeness, I include in the last three columns the outcomes from OLS estimation of Equation (5). Even using a different source of data for Internet diffusion, this is estimated to have a negative and significant impact on total (-1.6%) and property (-1.7%) crime rates. The IV estimated coefficients are considerably close in magnitude to my baseline results. In addition, these outcomes do not discredit my assumption that the effect on crime of broadband penetration in the period 2000-2012 approximates that of the Internet diffusion at large. In conclusion, this specification check further enhances confidence in my baseline IV estimation results.

6.1.4 Local level analysis

States are huge aggregates and a great deal of variation goes on within their boundaries. Because of this, a concern with my baseline analysis is that the estimation method may not be appropriate as too big entities are being considered. To address this issue, I replicate the analysis at a more disaggregated level. To the best of my knowledge, the only source of Internet data at local level is represented by the CPS Internet

and Computer Use Supplements, which report data at the Metropolitan Statistical Area (MSA) level. MSAs are geographic entities institutionally delineated for statistical purposes and consisting of one or more counties.¹² Caution shall be paid because CPS data are not representative for all parts of the US at the MSA level.¹³ Still, this is the only source available to perform a specification check at a disaggregated level. I focus on a variable indicating whether the household has an Internet connection, regardless of the connection technology – the survey question is: “Does anyone in this household connect to the Internet from home? Yes/No”. This variable is available in six years over the period 2000-2012 (namely: 2000, 2001, 2003, 2007, 2009 and 2010), and has been used in previous related literature (Kendall, 2007).

Local data on crime are available by FBI’s UCR at the county level. They include a coverage indicator (increasing from 0 to 100) which represents a measure of data quality for each county per year combination. I set a minimum threshold for acceptable data quality at 90/100 and replaced all entries that did not satisfy such requirement with missing values (I replicated the analysis using different floor thresholds to make sure results do not change significantly). Finally, I transformed county-level data into MSA-level data and merged them with the Internet diffusion data in a yearly (six years with gaps), MSA-level panel. MSAs are re-delineated periodically. I focus on the 112 MSAs for which data on the dependent and control variables are available in all sample years. This number corresponds to about one third of the total number of MSAs. Table 6 displays the mean and standard deviation for the MSA-level, time-varying variables described above in years 2000 and 2010, as well as their percentage variation in mean over such period. Data are in line with state-level figures presented in Table 1 (even though those statistics referred to years 2000 and 2012). Data confirm that crime rates have been falling while the Internet was expanding. Crime rates are a bit higher at the MSA level in comparison with the state level. This is likely to be due to the fact that MSAs are by definition densely populated areas.

For this local level study, I implement both an IV and an OLS setup that parallel the state level analysis. As instrumental variables, I use residential telephone adoption rate in 1986 and the share of population employed in the Cable TV sector in 1990, both at the MSA level. Telephone adoption data are available by the CPS at the MSA level, and 1986 is the first year in which information is reported for a considerable number of MSAs. Data on employment in the Cable TV segment at the MSA level is provided by the US Bureau of Labor Statistics.

Equation (1) and Equation (2) become, respectively:

$$(7) \quad y_{st} = \beta_0 + \beta_1 HHint_rate_{st} + \rho_t + \theta_s + \varepsilon_{st}$$

$$(8) \quad HHint_rate_{st} = \beta_0 + \beta_1(\rho_t * phonerate_s) + \beta_1(\rho_t * cableTV_s) + \rho_t + \theta_s + \varepsilon_{st}$$

Where the subscript s denotes MSA, $HHint_rate_{st}$ is the natural logarithm of the percentage of households that have an Internet connection and θ_s are MSA fixed effects. I use the robust matrix variance estimator with standard errors clustered at the MSA level.

In contrast with respect to the state level outcomes, my instruments do not seem to work well in the local setup. In fact, first stage results in Table 7 show that there is no significant association between the instruments and the endogenous variable in any of the sample years, implying a lack of relevance for the instruments at the local level.

The first three columns in Table 8 present the second stage results from the IV model, whereas the last three columns display the outcomes from OLS estimation of Equation (7). All estimated coefficients are negative and point estimates are not distant in magnitude from those obtained in my baseline analysis. However, possibly as a result of the lack of relevance of the instruments, only the OLS estimated coefficients for total and property crime are significant in this setting.

Although no considerations on causality of effects can be drawn from the local analysis, outcomes from this specification lend support to the finding of a negative and significant association between Internet diffusion and crime, in line with my state-level results.

6.2 Heterogeneity analyses

Results uniformly point to a negative and significant relationship between Internet diffusion and total and property crime. Relating to previous literature, these outcomes suggest that any content-driven

¹² See <http://www.census.gov/population/metro/>.

¹³ See <http://www.census.gov/cps/about/faq.html#Q3>

behavioural effect is dominated by the substitution effect. Previous research put forth that the latter may be heterogeneous across offense types and/or demographic groups. This section aims at testing these heterogeneity hypotheses.

6.2.1 Effect of broadband diffusion on specific crimes

The negative effect of broadband diffusion on crime may be more intense on that portion of delinquency that can be understood as recreational rather than predetermined and calculated. The rationale behind this stems directly from the very concept of substitution effect. Theory suggests that the negative net effect of Internet diffusion on crime reflects a modification in people's allocation choices, so that in terms of both time and energy online diversionary activities crowd out offline diversionary activities that more likely lead to crime. If this line of reasoning is correct, we shall observe that those types of offenses that are more likely to be committed in a recreational fashion have been more intensely affected by broadband penetration. The concept of recreational crime is usually associated with drug use and related offenses. Unfortunately, this is not one of the categories included in the crime data I am using. Then, the only hypothesis which seems reasonable to make is that especially severe offenses, like murder and, more generally, violent crimes are less likely to be pursued with a recreational intent – at least, this should be the case given that violent crimes are typically associated with more severe punishments. Then, if the suggested heterogeneity across offense types existed, data should display that property offenses have been affected more intensely with respect to violent offenses. My baseline results already lend some support to this hypothesis, in that a significant effect of Internet penetration is found on overall property crime and not on overall violent crime at large. A further step in this direction can be taken by replicating the baseline IV analysis using, as dependent variables, specific violent and property offenses.

Table 9 shows the findings. Remarkably, the effects on each property crime are significant. All else being equal, a broadband penetration increase by one additional line per 100 people is associated with a 1% - 2.3% decrease in burglary, larceny theft and motor vehicle theft rates. On the other hand, the estimated impact of broadband expansion on each of the violent offenses is not significantly different from zero. These results are consistent with the hypothesis that the diffusion of the Internet had a more intense impact on that portion of crime that can be understood as recreational, implying the substitution of one diversionary activity with another.

Previous empirical research studied the link between Internet and sex crimes. In particular, Bhuller *et al.* (2011) used data referring to Norway in an IV setup and found that Internet use is associated with a significant increase in sex crimes, likely as a result of a behavioural effect driven by the increased consumption of pornography. On the other hand, Kendall (2007) reached opposite conclusions on the association between Internet use and rape by studying US state-level data over the period 1998-2003. My analysis suggests yet another conclusion – that Internet diffusion does not significantly impact the occurrence of forcible rapes.

6.2.2 Effect of broadband diffusion on arrest rates

It has been suggested that the effect of Internet penetration may be heterogeneous across different demographic groups. In particular, Griffiths and Sutton (2013) consider that the substitution effect may be stronger on the youth, who typically contribute more than proportionally to delinquency and are particularly keen on new technologies. In the words of the authors, “all the time spent online must equate to less time on the street leading to less potential offending time and a smaller population of available victims of violence and robbery”. I aim at empirically testing this hypothesis. An important data limitation is that there is no direct measure of the number of crimes committed by age group. Only when a crime is cleared by an arrest is it possible to attach an age to the criminal. I employ FBI's UCR data on arrests by age and sex to and replicate my baseline IV analysis taking, as dependent variables, arrest rates for different combinations of age and gender. In particular, I consider the natural logarithm of the number of people arrested at age 0-14, 0-19, 0-24, 15-19, 15-24, 20+, 30+ as a percentage of people in that age range. I also consider the percentage of males only arrested at age 0-19, 15-19 and 15-24 – this is because males are arrested at a considerably higher rate than females.

Second stage results are shown in Table 10. The data does not seem to endorse the heterogeneous effect proposed by Griffiths and Sutton (2013). In fact, none of the estimated coefficients is significantly

different from zero. Therefore, I cannot detect a systematic variation in the effect of Internet diffusion across different demographic groups.

7 Discussion and economic significance

Results uniformly point to a negative and significant relationship between broadband Internet diffusion and total and property crime rates. Relating these outcomes to previous theoretical literature, a negative net impact of broadband Internet diffusion on crime suggests that any content-driven behavioural effect is dominated by the substitution effect – time and energy focused online crowd out alternative activities that are more likely to lead to crime. In addition, my results lend support to the notion of heterogeneous effects by type of offense, as broadband Internet diffusion seems to significantly affect property but not violent crime. In particular, I argue that this outcome fits the prevision that recreational offenses are being substituted by alternative forms of diversion. This is because I expect mostly property crimes rather than violent crimes to be committed with a diversionary intent – at least, this should be the case given that violent offenses are typically associated with more severe punishments. Linking my results to previous empirical research, the finding that the net effect of broadband diffusion on crime is negative is in line with the conclusions reached by those studies that analysed the effects on crime of technological platforms other than the Internet. These studies find evidence that both an arousal and a substitution effects exist, and that the latter more than outweighs the former – so that the net effect on crime is negative. The setup of the present research does not allow me to separate between arousal and substitution effects. However, the finding of a negative net effect is consistent with the notion that the most relevant mechanism through which the Internet, just like videogames and movies, affects crime is by making people substitute away from alternative activities that would more likely lead to delinquency.

It is important to put the size of my baseline estimates into perspective. My IV estimates indicate a 1.4% decrease in total crime rates as broadband penetration increases by one additional line per 100 people. Broadband diffusion went from 2 to 67 lines per 100 people between 2000 and 2012 (see Table 1). Then, the estimated variation in total crime rate due to broadband expansion (holding all other covariates fixed) in the period 2000-2012 can be calculated as:

$$(9) \quad \Delta \ln(y) = \beta \Delta x$$

This leads to a substantial effect on crime, of roughly -60%. This figure is huge, especially considering that the actual variation in total crime rate over the period was about -20.4%. Two factors may help explain the magnitude of the estimated impact. First, the calculation assumes that all other covariates are held constant while in reality they are not, so that the overall evolution of crime depends on the combination of a variety of factors pressuring delinquency both upwards and downwards. Secondly, my IV estimates must be interpreted as a LATE. The source of exogenous variation in my IV setup comes from the deployment of the telephone and Cable TV networks in early years. This allows me to identify the effect of broadband internet deployment on crime rates for the sub-population ('compliers') for which broadband penetrated according to the patterns of pre-existing telephone and Cable TV networks. Then, if some compliers' characteristics relevant to crime evolution are not homogenous with respect to the overall population of users, the LATE will represent a biased estimate of the effect of Internet diffusion on the population at large. In the case at issue, since DSL and Cable Modem broadband technologies were by far the most widely spread until the end of the 2000s, I can expect the bias to be small. In addition, it is reasonable to think that such bias would be upwards, so that the LATE actually overstates the impact of the Internet diffusion on crime for the population at large. The reasoning behind this goes as follows. By construction, the estimated LATE is plausibly capturing the effect of fixed broadband Internet connections. In fact it is sound to believe that compliers have a DSL or Cable Modem connection – exactly because these technologies rely on the pre-existing telephone and Cable TV networks. Both of these are fixed residential technologies, and this implies that an individual must be at home to use them. On the other hand, the non-complier users are likely to mostly rely on a Mobile Wireless connection. This is because the latter (i) does not rely on the pre-existing telephone and Cable TV networks, so it is not captured by the instruments; and (ii) over the period 2000-2012, it has represented the third most widespread residential broadband platform after DSL and Cable Modem (see Figure 1). Then, relating to the discussion about the substitution mechanisms in Section 2.1, it seems reasonable to expect that the

effect is more intense on compliers – who surf the Internet from home – rather than on the rest of users – who access the Internet from their mobiles. This is because, while a content-drive catharsis effect may be at play for all users, the voluntary incapacitation mechanism – by which Internet negatively affect crime by simply keeping people off the streets – only affects the compliers. Because of this underlying difference between the sub-populations, the estimated LATE is likely to be higher than the effect for the population at large. In this context, the OLS estimates may be of some help. As already mentioned, these coefficients are likely to suffer from an omitted variable bias. However, they are not biased towards a specific sub-population and in this sense more closely reflect the effects of Internet diffusion on crime for the population at large. The OLS estimated coefficients indicate that, all else being equal, a broadband penetration increase by one additional line per 100 people is associated with a 0.6% decrease in total crime rates, implying an estimated variation in total crime rate due to broadband expansion about half the magnitude of the 2SLS estimated effect. This suggests that the LATE may indeed be overstating the magnitude of the impact of Internet diffusion on crime. Importantly, though, the ongoing considerations do not notch the core result of the analysis – that the overall effect of Internet diffusion on crime is negative, significant and substantial.

I have performed a number of falsification and specification checks to enhance confidence in my outcomes. Overall, these analyses lend support to validity of my IV approach and show that results are robust with respect to the chosen source of Internet data and level of disaggregation.

Further insights from the analysis suggest that the diffusion of the Internet has a more intense impact on that portion of crime that can be understood as recreational, implying the substitution of one diversionary activity with another. On the other hand, I find no evidence that the effects of Internet diffusion on crime are heterogeneous across demographic groups. Still, this piece of analysis suffers from the limitation that there is no direct measure of the number of crimes committed by age group.

8 Conclusions

How does Internet diffusion affect crime? Theoretical studies have pointed to both a positive behavioural impact and a negative substitution effect. The empirical literature on the effects of Internet expansion on delinquency is scarce and mainly focused on a specific kind of offenses – sex crimes. Non-Internet-specific empirical research has been carried on in related fields, providing support to the existence of both the behavioural and the substitution effects. Thus, the overall impact of the Internet diffusion on the most common types of crime is *a priori* unclear. This paper aims at casting light on the issue.

I exploit the variation in broadband penetration and criminal activity in the US over the period 2000-2012. To avoid the omitted variable bias which is likely to affect the OLS estimates, I resort to an IV approach. My underlying assumption is that the deployment of residential telephone and Cable TV networks in early years provides a relevant source of exogenous variation in recent broadband diffusion. This is because the two most widespread non-mobile residential broadband platforms in the US rely on the pre-existing physical networks of these communication technologies. I investigate this hypothesis through a number of tests, and outcomes lend support to the validity of this approach. My IV results suggest that the overall impact of Internet diffusion on the most common property crimes is negative, significant and substantial. On the other hand, violent crimes seem not to be affected. A number of tests enhance confidence in the plausibility of a causal interpretation of my results.

This study speaks to the ongoing debate as to the far-reaching social impacts of the Internet, and whether and how data traffic should be restricted or monitored to prevent potential adverse effects. Overall, my findings advise against these actions. Still, caution is in order. From a policy perspective, it is important to disentangle the behavioural and substitution effects. In cases where the former is relevant (although more than outweighed by the latter), monitoring or restricting some types of contents for some ranges of users may make the overall negative effect even more substantial.

In addition, this research contributes to the stream of literature investigating the crime drop experienced by the US since the early 1990s. My findings suggest that Internet diffusion should be enlisted among the factors that have jointly concurred to sustaining the decline.

References

- Anderson, C.A., Gentile, D.A. and Buckley, K.E. (2007) *Violent video game effects on children and adolescents: theory, research and public policy*. Oxford University Press, 1st edition.
- Anderson C.A., Shibuya, A., Ihori, N., Swing, E.L., Bushman, B.J., Sakamoto, A., Rothstein, H.R. and Saleem, M. (2010) "Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review." *Psychological Bulletin*, Vol. 136(2): 151-73.
- Angrist, J. and Pischke, S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.
- Bellou, A. (2015) "The Impact of Internet Diffusion on Marriage Rates: Evidence from the Broadband Market." *Journal of Population Economics*, Vol. 28: 265–297.
- Bhuller, M., Havnes, T., Leuven, E. and Mogstad, M. (2013) "Broadband Internet: An Information Superhighway to Sex Crime?." Review of Economic Studies, Oxford University Press, Vol. 80(4): 1237–1266.
- Cashmore, J. (2012) "Crime Reducing Entertainment: The Contribution Of Media Entertainment And Communication Technologies To The UK's Victimisation Drop." *Internet Journal of Criminology*, 2012. Available [here](#).
- Copus, R. e Laqueur, H. (2014) "Entertainment as Crime Prevention: Evidence from Chicago Sports Games" Available [here](#).
- Cunningham, S., Engelstätter, B., and Ward, M. (2011) "Understanding the effect of violent video games on violent crime." ZEW - Centre for European Economic Research Discussion Paper No. 11-042. Working Paper Series.
- Dahl, G. and Della Vigna, S. (2009) "Does movie violence increase violent crime?" *The Quarterly Journal of Economics*, Vol. 124(2): 677–734.
- Dilulio, J. (1996) "Help Wanted: Economists, Crime, and Public Policy." *Journal of Economic Perspectives*, Vol. 10: 3–24.
- Donohue, J. e Levitt, S. (2001) "Legalized Abortion and Crime." *Quarterly Journal of Economics*, Vol. 116(2): 379–420.
- Duflo, E. (2001) "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment." *American Economic Review*, Vol. 91(4): 795–813.
- FCC "The Broadband Availability Gap, 2010". Available [here](#).
- FBI – Internet Crime Complaint Center (IC3) "2013 Internet crime report". Available [here](#).
- FBI – Uniform Crime Reports (UCR) "Crime in the United States 2013". Available [here](#).
- Fox, J. (1996) "*Trends in Juvenile Violence: A Report to the United States Attorney General on Current and Future Rates of Juvenile Offending*." Washington, D.C.: Bureau of Justice Statistics.
- Griffiths, M. and Sutton, M. (2013) "Proposing the Crime Substitution Hypothesis: Exploring the possible causal relationship between excessive adolescent video game playing, social networking and crime reduction." *Education and Health Journal*, Vol. 31(1): 17–21.
- Imbens, G.W. and Angrist, J.D. (1994) "Identification and estimation of local average treatment effects." *Econometrica*, Vol. 62(2): 467–75.
- ITU (2012) "The Impact of Broadband on the Economy: Research to Date and Policy Issues." Available [here](#).
- Kelly, M. (2000) "Inequality And Crime." *The Review of Economics and Statistics*, Vol. 82(4): 530–539.
- Kendall, T. D. (2007). "Pornography, rape, and the Internet." Mimeo, Clemson University.
- Kuhn, P. and Skuterund, M. (2004) "Internet job search and unemployment durations." *American Economic Review*, Vol. 94(1): 218–232.
- Kuhn, P. and Mansour, H. (2014) "Is Internet Job Search Still Ineffective?" *The Economic Journal*, Vol. 124: 1213–1233.
- Lehr, W.H., Osorio, C.A., Gillett, S.E. and Sirbu, M.A. (2006). "Measuring broadband's economic impact." Final Report. National Technical Assistance, Training, Research and Evaluation Project # 99-07-13829. Available [here](#).
- Levitt, S.D. (2004) "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not." *Journal of Economic Perspectives*, Vol. 18(1): 163–190.
- King, M., Ruggles, S., and Trent, A., Flood, S., Genadek, K., Schroeder ,M. B., Trampe, B., and Vick, R. (2010) Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota.
- NTIA (2004) A nation online: entering the broadband age. U.S. Department of Commerce, Economics and Statistics Administration.
- NTIA (2011) Exploring the digital nation: home broadband Internet adoption in the United States. U.S. Department of Commerce, Economics and Statistics Administration.
- NTIA (2013) Household Broadband Adoption Climbs to 72.4 Percent. Available [here](#).
- Stevenson, B. (2006) "The impact of the Internet on worker flows". Working paper. The Wharton School, University of Pennsylvania.
- Stock, J.H., Wright, J. H. and Yogo, M. (2002) A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments, *Journal of Business and Economic Statistics*, 20(4), 518–29.
- Sutton, M. (2010) "Criminology: The Blog of Mike Sutton. Routine Activities Theory, The Internet and the 15-Year Crime Drop." Available [here](#).
- Sutton, M. (2011) "The Internet and the unfathomable crime drop." *Dysology: The Study of Remarkable Claptrap*. [Here](#).
- Wooldridge, J. M. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, Mass., 2002.
- Zimring, F.E. and Hawkins, G. *Crime Is Not the Problem: Lethal Violence in America*. New York: Oxford University Press, 1997.
- Zimring, F.E. *The great American crime decline*. New York: Oxford University Press, 2007.

Figures and tables

Figure 1: Residential broadband Internet lines by technology (%), 2000-2012

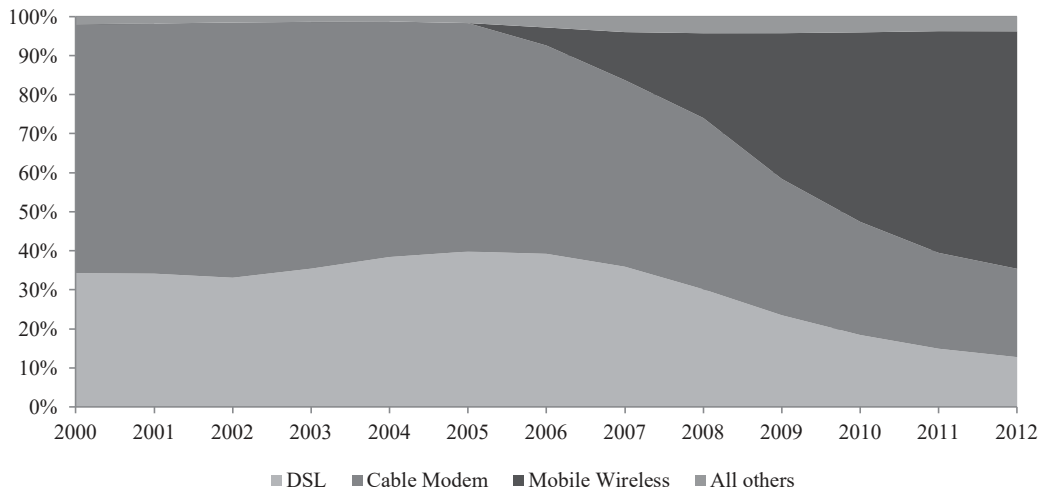


Figure 2: Residential broadband lines (per 100 people): actual vs predicted by rate of residential telephone adoption in 1955 and rate of employment in the Cable TV sector in 1990

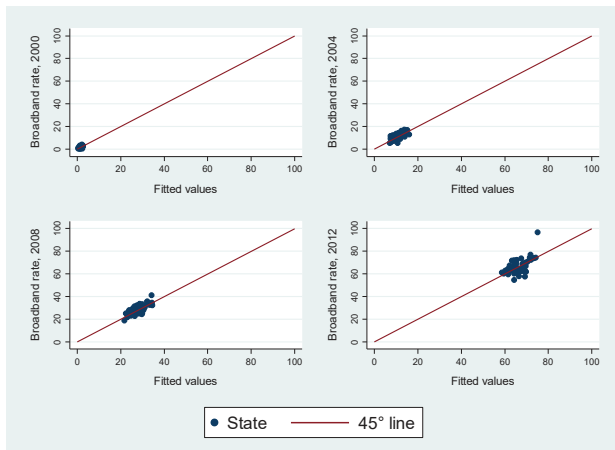
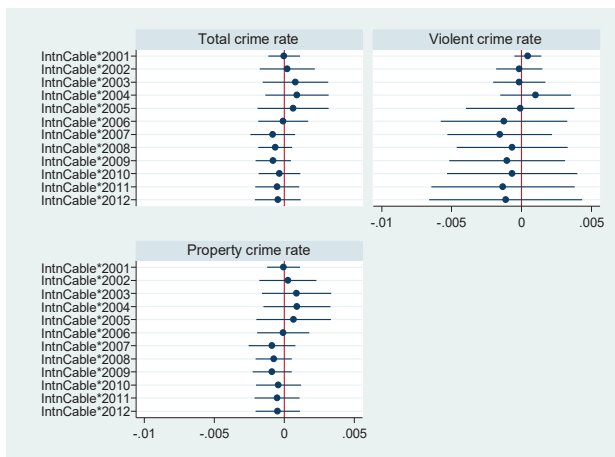
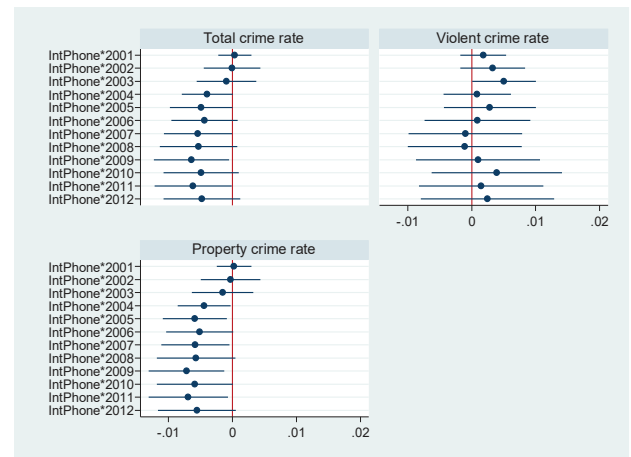


Figure 4: Reduced form estimation, $\rho_t * cableTV_s$



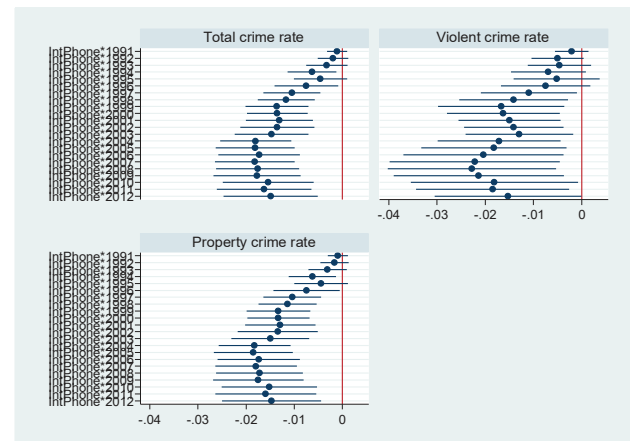
Note: The estimated coefficients are displayed with their respective 95% confidence intervals.

Figure 3: Reduced form estimation, $\rho_t * phonerate_s$



Note: The estimated coefficients are displayed with their respective 95% confidence intervals.

Figure 5: Falsification test, $\rho_t * phonerate_s$, 1990-2012



Note: The estimated coefficients are displayed with their respective 95% confidence intervals.

Table 1: Descriptive statistics, 2000 and 2012

Variable	2000		2012		Change (%)
	Mean	SD	Mean	SD	
Total crime rate	4,052.73	1,040.01	3,225.61	737.87	-20.41
Violent crime rate	443.33	241.14	371.27	178.66	-16.25
Property crime rate	3,609.40	868.12	2,854.33	606.76	-20.92
Residential broadband lines per 100 people – FCC	1.58	0.94	66.95	6.77	4,137.34
Households with Internet connection (%) – CPS	45.11	6.70	73.96 ^(*)	6.11 ^(*)	64
Black or Afro-Americans (%)	11.15	11.91	11.54	10.99	3.50
Male population aged 15-24 (%)	7.27	0.64	7.17	0.39	-1.38
Population density (/sqmi)	315.78	1,172.75	345.24	1,294.97	9.33
Personal income per capita (\$)	29,441.35	4,804.47	44,014.27	7,956.04	49.50
Unemployment rate (%)	3.91	0.93	7.37	1.71	88.49

Note: Crime rates are per 100,000 people. ^(*) These values refer to year 2010.

Table 2: First stage estimation

DV: Residential broadband lines rate			
Interaction	*Phone adoption rate	*Rate of employment in Cable TV sector	
(year==2001)*	0.0217 [0.0209]	0.0241* [0.0127]	
(year==2002)*	0.044 [0.0344]	0.0364*** [0.0138]	
(year==2003)*	0.1551*** [0.0578]	0.0525*** [0.0186]	
(year==2004)*	0.2083*** [0.0630]	0.0665*** [0.0247]	
(year==2005)*	0.2848*** [0.0811]	0.0600** [0.0236]	
(year==2006)*	0.2885*** [0.0672]	0.0491* [0.0277]	
(year==2007)*	0.3023*** [0.0715]	0.0886*** [0.0183]	
(year==2008)*	0.3103*** [0.0734]	0.0967*** [0.0142]	
(year==2009)*	0.2699*** [0.0667]	0.1240*** [0.0149]	
(year==2010)*	0.2085** [0.0846]	0.1084*** [0.0235]	
(year==2011)*	0.2323* [0.1204]	0.0851** [0.0359]	
(year==2012)*	0.1905* [0.0993]	0.1222*** [0.0362]	
Year effects	Yes		
Further controls	Yes		
N	636		
F-test for IV	9.7528		
F-test for IV - Prob > F	0		

Note: Robust standard errors (in brackets) are clustered at the state level. Further controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.

Table 3: Second stage and OLS estimation

DV: Crime rate	2SLS			OLS		
	Total	Violent	Property	Total	Violent	Property
Broadband lines rate	-0.0141*** [0.0051]	-0.0094 [0.0114]	-0.0153*** [0.0051]	-0.0060** [0.0026]	-0.0032 [0.0040]	-0.0065** [0.0027]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Further controls	Yes	Yes	Yes	Yes	Yes	Yes
N	636	636	636	656	656	656
Hansen J statistic	24.757	27.982	23.777			
p-value	0.3629	0.2165	0.4163			

Note: Robust standard errors (in brackets) are clustered at the state level. Further controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.

Table 4: Second stage estimation – Robustness test with time interactions and LIML

DV: Crime rate	Time trend (A)			Time dummies (B)			LIML (C)		
	Total	Violent	Property	Total	Violent	Property	Total	Violent	Property
Broadband lines rate	-0.0154*** [0.0051]	-0.0217** [0.0106]	-0.0153*** [0.0051]	-0.0204*** [0.0054]	-0.0200** [0.0094]	-0.0205*** [0.0057]	-0.0155*** [0.0059]	-0.01 [0.0124]	-0.0170*** [0.0059]
Year effects	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Further controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	636	636	636	636	636	636	636	636	636

Note: Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively. For column sets (A) and (B): Further controls include contemporaneous exogenous covariates and interactions between baseline year (2000) exogenous covariates and a linear time trend (column set A) or time dummies (column set B). Contemporaneous exogenous covariates are population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. I use the 2SLS estimator. For column set (C): Further controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24.

Table 5: Second stage and OLS estimation – CPS Internet data

DV: Crime rate	2SLS			OLS		
	Total	Violent	Property	Total	Violent	Property
HH has Internet (%)	-0.0157** [0.0075]	-0.0092 [0.0089]	-0.0170** [0.0077]	-0.0041** [0.0019]	-0.0029 [0.0027]	-0.0043** [0.0020]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Further controls	Yes	Yes	Yes	Yes	Yes	Yes
N	294	294	294	306	306	306

Note: Robust standard errors (in brackets) are clustered at the state level. Further controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.

Table 6: Descriptive statistics – MSA-level data, 2000 and 2010

Variable	2000		2010		Change (%)
	Mean	SD	Mean	SD	
Total crime rate	4357.87	2023.63	3745.89	1979.23	-14.04
Violent crime rate	453.05	262.11	417.92	271.81	-7.75
Property crime rate	3904.83	1813.73	3327.97	1745.6	-14.77
CPS – HH with Internet connection (%)	47.94	10.76	74.95	9.58	56.34

Note: Crime rates are per 100,000 people.

Table 7: First stage estimation – MSA-level analysis

DV: Household has Internet connection (%)			
Interaction	*Phone adoption rate	*Rate of employment in Cable TV sector	
(year==2001)*	1.871	-0.0359	
	[5.4539]	[0.0371]	
(year==2003)*	6.8351	-0.0406	
	[5.6994]	[0.0382]	
(year==2007)*	-4.361	-0.0249	
	[4.9199]	[0.0379]	
(year==2009)*	4.6845	-0.0074	
	[4.9342]	[0.0365]	
(year==2010)*	-2.5501	-0.0343	
	[6.2585]	[0.0398]	
Year effects		Yes	
Further controls		No	
N		540	
F-test for IV		1.16	
F-test for IV - Prob > F	0.2382		

Note: Robust standard errors (in brackets) are clustered at the MSA level. No further controls are included. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.

Table 8: Second stage and OLS estimation – MSA-level analysis

DV: Crime rate	2SLS			OLS		
	Total	Violent	Property	Total	Violent	Property
HH has Internet (%)	-0.0118 [0.0120]	-0.0057 [0.0137]	-0.0126 [0.0120]	-0.0046** [0.0019]	-0.0034* [0.0020]	-0.0047** [0.0020]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Further controls	No	No	No	No	No	No
N	531	531	531	658	658	658

Note: Robust standard errors (in brackets) are clustered at the MSA level. No further controls are included. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.

Table 9: Second stage estimation – Specific crimes

DV: Crime rate	Violent crimes				Property crimes		
	Murder	Forcible rape	Robbery	Aggravated assault	Burglary	Larceny theft	Motor vehicle theft
Broadband lines rate	0.0026 [0.0130]	-0.0121 [0.0106]	-0.002 [0.0088]	-0.0129 [0.0149]	-0.0235*** [0.0074]	-0.0100** [0.0050]	-0.0231** [0.0101]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	636	636	636	636	636	636	636

Note: Robust standard errors (in brackets) are clustered at the state level. Further controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. I use the 2SLS estimator. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.

Table 10: Second stage estimation – Arrests rates by demographic group

DV: Arrest rate	Age 0-14	Age 0-19	Age 0-24	Age 15-19	Age 15-24	Age 20+	Age 25+	Age 30+	Age 0-19 Males	Age 15-19 Males	Age 15-24 Males
Broadband lines rate	-0.012 [0.0195]	0 [0.0159]	-0.0014 [0.0174]	0.0017 [0.0165]	0.0004 [0.0176]	-0.0105 [0.0214]	-0.0143 [0.0225]	-0.0155 [0.0228]	-0.0009 [0.0162]	0.0007 [0.0166]	0.0003 [0.0179]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	621	621	621	621	621	621	621	621	621	621	621

Note: Robust standard errors (in brackets) are clustered at the state level. Further controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. I use the 2SLS estimator. The estimation was replicated without the latter control: results do not change significantly. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.

MOBILIDADE DOS FATORES PRIMÁRIOS E SEUS EFEITOS SOBRE O CRESCIMENTO ECONÔMICO E O BEM-ESTAR GERADOS PELO CRÉDITO RURAL NAS REGIÕES BRASILEIRAS

Talita Priscila Pinto^a Erly Cardoso Teixeira^b Ângelo Costa Gurgel^c

^aDepartamento de Economia Aplicada, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil.

^bDepartamento de Economia Aplicada, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil.

^cEscola de Economia de São Paulo, Fundação Getúlio Vargas, São Paul, Brasil.

Área 3: Microeconomia Aplicada

RESUMO

O presente trabalho investiga os efeitos da mobilidade dos fatores produtivos sobre o crescimento econômico e o bem estar social gerados pela política de Equalização das Taxas de Juros (ETJ) ao setor agrícola nas regiões brasileiras. Utiliza-se do Projeto de Análise de Equilíbrio Geral da Economia Brasileira (PAEG) nas simulações analíticas. É analisado um cenário em que o valor do subsídio da política de ETJ e do crédito rural subsidiado disponibilizado pela ETJ são eliminados da agropecuária, sendo que o crédito subsidiado é realocado entre os diversos setores da economia. O cenário é analisado considerando três diferentes situações de mobilidade dos fatores produtivos: ausência de mobilidade, mobilidade parcial e mobilidade perfeita. Todos os resultados da pesquisa são apresentados com os sinais trocados a fim de se obter os efeitos da política de ETJ e do crédito gerado por essa política na economia. Os resultados revelam crescimento do PIB nas regiões brasileiras menor que o custo da política. Portanto, a política de ETJ, em termos de geração de crescimento econômico, apresenta taxa de retorno negativa. Quando se considera a análise em termos de bem-estar, os efeitos do choque são positivos e maiores que o custo da política para todas as regiões brasileiras. Portanto, a política de ETJ, em termos de geração de bem-estar, apresenta taxa de retorno positiva. Palavras chave: crédito rural, mobilidade de fatores, equilíbrio geral, PAEG.

ABSTRACT

This paper investigates the effects of primary production factor mobility on economic growth and welfare generated by the interest rate equalization policy (ETJ) in the agricultural sector in the Brazilian regions. This study uses the General Equilibrium Analysis Project for the Brazilian Economy (PAEG) to perform the analytical simulations. The study analyzes a scenario in which the value of the ETJ policy and the subsidized rural credit provided by the ETJ are eliminated from agriculture. The subsidized credit is reallocated among the various sectors in the economy. The scenario is analyzed considering the mobility of three different primary production factors among Brazilian regions: zero mobility, partial mobility and complete mobility. The results are presented with the signals exchanged to obtain the effects of ETJ policy on the economy. The results suggest that GDP growth is lower than the subsidy cost in all Brazilian regions except in the Midwestern and Southern regions with complete factor mobility. In terms of generating economic growth, the ETJ policy therefore presents a negative rate of return. When one considers the analysis in terms of welfare, the shock effects are positive and higher than the cost of the policy to all regions of Brazil. In terms of welfare generation, the ETJ policy therefore presents a positive rate of return.

Keywords: rural credit, primary factors mobility, general equilibrium, PAEG

Classificação JEL: C68 / Q14 / Q18

1. Introdução

Políticas intervencionistas em mercados agrícolas são eventos comuns em países desenvolvidos e em desenvolvimento. Entretanto são criticadas por diversas instituições multilaterais como o Banco Mundial, o Fundo Monetário Internacional (FMI) e a Organization for Economic Co-operation and Development (OECD). É postulado pela teoria clássica, nos modelos de equilíbrio parcial que os subsídios agrícolas geram ineficiências alocativas, distributivas, além de custo social. A maioria dos países desenvolvidos, entretanto, adota uma política intervencionista insistindo na prática de subsídios. A justificativa desses governos é que a ausência de proteção faria com que muitos produtores abandonassem suas atividades, agravando problemas sociais. A partir do exposto questiona-se se as justificativas sociais são o agente motivador da prática de subsídios ou se essa política promove crescimento econômico maior que seu custo.

O lado social engloba a manutenção do emprego no campo, o poder de compra da população rural, entre outros argumentos, utilizados como alicerce para adoção das políticas agrícolas citadas, mas, além das razões sociais, a manutenção dos subsídios pode ocorrer por razões econômicas. Gasques e Villa Verde (2003), Castro e Teixeira (2004), Cardoso, et.al (2011), consideram, além do lado social, a existência de resultados positivos sobre o crescimento econômico brasileiro.

Um estudo de Taylor (1994), diz que incentivos agrícolas combinados com elevação de renda, podem gerar um processo de desenvolvimento econômico já que também impactam positivamente setores não agrícolas.

Um estudo feito por Cardoso et.al (2014), com dados de 2004, simula, utilizando um modelo de equilíbrio geral, o PAEG, a eliminação do subsídio agrícola recebido por meio da equalização da taxa de juros e a eliminação do consumo de insumos intermediários proporcionado pelo crédito rural subsidiado a esse setor. Dessa forma, a autora estabelece a importância da política ETJ com a promoção do crescimento econômico e do bem-estar nas regiões brasileiras.

Segundo Mundell (1961), o equilíbrio geral da economia é afetado pela mobilidade dos fatores de produção. A teoria neoclássica da produção postula que há migração de mão-de-obra das regiões de baixos salários para regiões com mais altos salários até que as diferenças sejam eliminadas. Em um mercado perfeito, o capital fluirá das regiões de baixo rendimento para as regiões de maior rendimento, até que novamente, as diferenças se anulem.

Segundo Souza (1981), a mobilidade dos fatores não é total, e sim dinâmica, e há uma série de restrições que ocorrem influenciando esse processo. Apesar da migração da mão-de-obra, despesas de viagem, distância e custo de instalação para o emigrante são alguns dos fatores limitantes. Em relação às restrições à mobilidade do capital, o autor considera a instabilidade das demandas regionais, imobilidade do capital devido a investimentos físicos como equipamentos ou infraestrutura em uma dada localização, além do fator segurança que dificulta o acesso de pequenas firmas ao capital. Ainda segundo Souza (1981), os elementos não são distribuídos uniformemente no território, existe heterogeneidade e descontinuidade, ausência de vias de transporte em todas as direções e concentração demográfica e industrial.

Este trabalho divide-se em mais três seções além desta introdução. Na segunda seção está a metodologia com a apresentação do modelo PAEG, a fonte de dados e os cenários analíticos. Na terceira seção, expõem-se os resultados. Na quarta seção, encontram-se as principais conclusões.

2. Metodologia

Modelos Aplicados de Equilíbrio Geral seguem uma base teórica walrasiana onde a economia é concorrencial e possui dois agentes principais, produtores e consumidores. Os agentes produzem, consomem e comercializam bens e fatores. Os consumidores, com suas restrições orçamentárias e cestas de preferências, demandam bens maximizando sua função utilidade. As preferências são, hipoteticamente, contínuas e convexas, e delas resultam funções de demanda contínuas e homogêneas de grau zero em relação aos preços, ou seja, somente os preços relativos podem ser determinados.

Do lado da produção, a tecnologia é descrita por uma função de produção com rendimentos constantes de escala, significando que, no equilíbrio, o lucro das firmas é nulo. As firmas são dotadas de uma determinada tecnologia de produção e demandam fatores de forma a minimizar seus custos. Esses modelos possibilitam a análise de efeitos diretos e indiretos advindos de alterações em políticas públicas, tais como choques tarifários, modificações em alíquotas de impostos e subsídios (TEIXEIRA, PEREIRA, GURGEL, 2013, p.14).

Para captar os efeitos alocativos e distributivos que uma política intervencionista pode gerar dentro de mercados agrícolas, as análises aplicadas de Equilíbrio Geral Computáveis são as mais indicadas, já que permitem que essa captação seja feita tanto para o mercado de bens quanto o de fatores e sobre a distribuição setorial da renda. Portanto o referencial teórico que corrobora a pesquisa em questão, toma como base a análise clássica de equilíbrio geral da economia.

2.1. O Modelo PAEG

O PAEG (Teixeira, Gurgel e Pereira, 2013) é um modelo estático, multiregional e multissetorial e teve sua elaboração baseada no *GTAPinGAMS* (Rutherford e Paltsev, 2000; Rutherford, 2005) que, por sua vez origina-se do *GTAP*¹ (Hertel, 1997; GTAP, 2001). Existem algumas diferenças entre os dois modelos. Diferentemente do GTAP, que utiliza a linguagem GEMPACK (Cotsi e Pearson, 1988), o PAEG adota a estrutura básica do modelo *GTAPinGAMS*, que foi elaborado como um problema de complementariedade mista não-linear, em linguagem de programação GAMS² (*General Algebraic Modeling System*, Brooke et al., 1998). Adicionalmente, segundo Teixeira, Gurgel e Pereira (2013) no PAEG a base de dados³ referente à economia brasileira foi desagregada a fim de representar suas cinco grandes regiões (Centro Oeste, Norte, Nordeste, Sul e Sudeste), mantendo intactos os dados do GTAP para as demais regiões do mundo, e os dados de fluxos comerciais entre o Brasil e as demais regiões do mundo.

O PAEG representa a forma como os bens e serviços são produzidos na economia brasileira e mundial. As regiões são representadas por uma estrutura de demanda final e o comportamento dos agentes é otimizador, eles maximizam seu bem-estar sujeitos à sua restrição orçamentária, considerando fixos o investimento e a produção do setor público. Os setores produtivos minimizam os custos com uma combinação de insumos intermediários e fatores primários de produção, dada a tecnologia. Os fluxos bilaterais de comércio entre as regiões, os custos de transporte, impostos e/ou subsídios também estão presentes na base de dados (GURGEL et.al, 2011). A Tabela 1 descreve os índices representados no modelo.

¹ Modelo de equilíbrio geral computável multiregional, multissetorial. Ver detalhes em www.gtap.org

² Sistema Geral de Modelagem Algébrica

³ Maiores informações sobre a conciliação dos dados das matrizes regionais brasileiras estão disponíveis em www.paeg.ufv.br, seção publicações, *Technical Papers* nº 1, 2 e 3.

Tabela 1: Índices de conjuntos da base de dados do modelo PAEG, 2007.

Índice	Descrição
i, j	Setores e bens
r, s	Países e regiões
f ∈ m	Fatores de produção de mobilidade livre dentro de dada região: trabalho e capital
f ∈ s	Fatores de produção fixos: terra e outros recursos naturais

Fonte: Gurgel *et al.*(2011).

O funcionamento do modelo pode ser demonstrado a partir das identidades contábeis macroeconômicas. A produção doméstica é distribuída entre exportações, serviços de transporte internacional, demanda intermediária, consumo privado, investimento, e consumo do governo. Bens importados são utilizados no consumo intermediário, no consumo privado e no consumo do governo.

Na produção do bem j (Y_{ir}) incluem-se insumos intermediários (domésticos e importados), fatores de produção móveis e consumo do agente público. A renda dos fatores de produção é distribuída ao agente representativo. O equilíbrio nos mercados de fatores é dado por uma identidade que relaciona o valor do pagamento dos fatores com a renda destes.

O equilíbrio entre oferta e demanda requer que as exportações sejam iguais às importações. Da mesma forma, a oferta agregada do serviço de transporte, é igual ao valor dos serviços de transporte nas exportações. O equilíbrio entre oferta e demanda, no mercado de serviços de transporte, iguala a oferta desses serviços à soma dos fluxos bilaterais de serviços de transporte adquiridos nas importações de bens.

A renda do governo é dada pela soma dos impostos e transferências. Assim a restrição orçamentária do governo pode ser representada pela equação (1).

$$vgm_r = \sum_i R_{ir}^Y + R_r^C + R_r^G + \sum_i R_{ir}^M + R_r^{HH} + vb_r \quad (1)$$

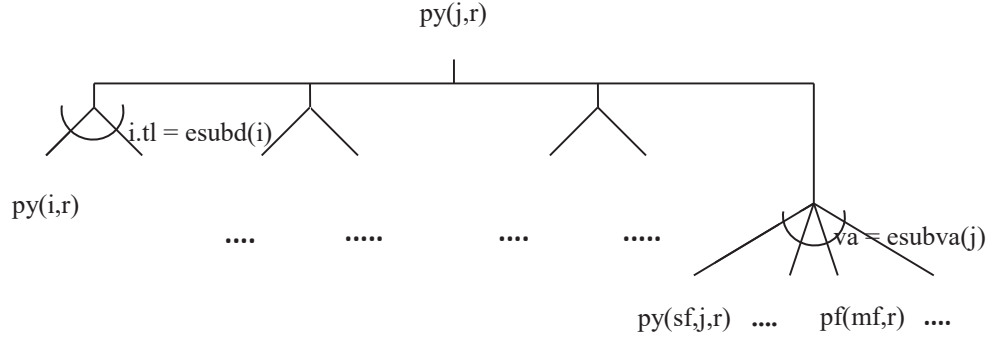
Onde (R_{ir}^Y), (R_r^C), (R_r^G) (R_{ir}^M) são impostos indiretos na produção e exportação, sobre consumo, na demanda do governo e nas importações, respectivamente. R_r^{HH} representa os impostos indiretos ao agente representativo, bem como transferências do exterior, vb_r .

A restrição orçamentária do agente representativo relaciona a renda dos fatores de produção, descontada dos pagamentos de impostos, com as despesas de consumo e investimento privado.

Portanto, partindo das identidades apresentadas, é possível visualizar dois tipos de condição: o equilíbrio de mercado (oferta igual à demanda para todos os bens e fatores de produção) e o balanço da renda (renda líquida igual à despesa líquida). No modelo PAEG, assim como no GTAP, consideram-se competição perfeita e retornos constantes à escala, de forma que os custos de produção se igualem ao valor da produção, e os lucros econômicos, a zero. Tal condição se aplica a cada um dos setores produtivos e atividades.

As identidades econômicas do modelo, contudo, não descrevem o comportamento dos agentes econômicos. Para entender o funcionamento do modelo, é preciso descrever como os agentes e setores se comportam. No entanto, nem todas as funções de comportamento serão apresentadas, para não ocorrer uma fuga ao objetivo do presente estudo.

O comportamento das firmas é otimizador e definido por funções de produção, e é representado em blocos de produção, uma vez que se utiliza a *syntax* do algoritmo MPSGE, desenvolvido por Rutherford (1999). Os setores produtivos combinam insumos intermediários e fatores primários de produção, a fim de minimizar custos, dada a tecnologia. Apresenta-se na Figura 1 a “árvore tecnológica” que representa o bloco de oferta de Y_{ir} e descreve as tecnologias assumidas pelas firmas nas indústrias do modelo.



$esubva(j)$: elasticidade de substituição entre os fatores de produção que compõem o valor adicionado.

$esubd(i)$: elasticidade de substituição entre os fatores domésticos e importados.

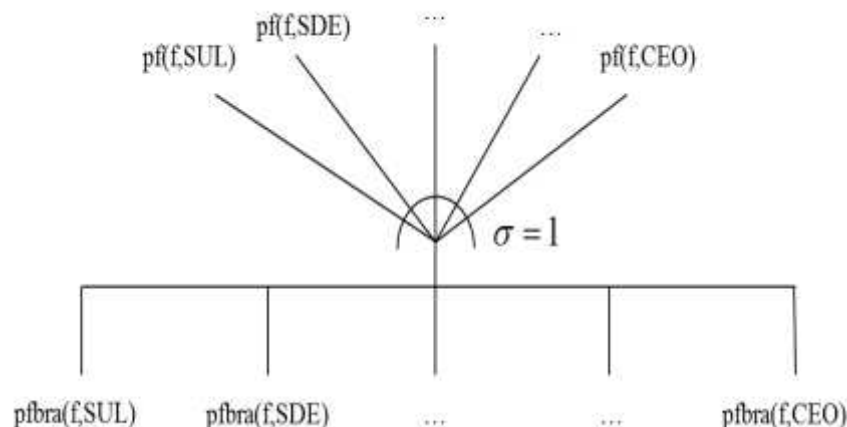
Fonte: Adaptado de Gurgel et al. (2013).

Figura 1: Árvore tecnológica da estrutura produtiva do PAEG.

Observa-se na Figura 1 que a oferta das firmas é definida por um problema de otimização e objetiva-se a minimização dos custos unitários, a partir da combinação de insumos primários de produção e insumos intermediários, domésticos e importados. Assim, primeiramente, as firmas decidem a combinação de fatores primários que será empregada ($py(sf, j, r)$) e ($pf(mf, r)$). A decisão é tomada com base na elasticidade de substituição entre os fatores da produção que compõem o valor adicionado ($esubva(j)$). Posteriormente, adquirem cestas de insumos intermediários, sobre as quais decidem entre bens domésticos e importados ($py(i, r)$) por meio da elasticidade de substituição ($esubd(i)$). O produto final é representado por ($py(j, r)$).

A Figura 2 representa o bloco de produção (ftr) responsável pela alocação de fatores entre regiões diferentes, em resposta a mudanças na economia. Nesse bloco, as dotações de um tipo de fator (f) advindas de todas as regiões estão disponíveis como insumos, para serem transformadas em fatores regionais que serão utilizados especificamente em cada uma das regiões. O símbolo σ determina a elasticidade de transformação de um fator de uma região em relação a esse mesmo fator de outra região.

Os fatores primários apresentados tem elasticidade de substituição igual à zero, ou seja, Leontief no equilíbrio inicial. Essas elasticidades definem que os fatores das diferentes regiões são sempre combinados em proporções fixas, de acordo com a dotação regional inicial. Após algum choque, esses fatores são distribuídos para as diversas regiões considerando uma função Cobb-Douglas de transformação entre as regiões, ou seja, não é possível o livre movimento do capital ou trabalho de uma região para outra, diante de diferenças nas remunerações dos fatores, uma vez que as características e composição dos fatores de cada região não são exatamente iguais.



σ = Elasticidade de transformação entre os fatores das diferentes regiões.

pfbra: Preço da dotação de fator de cada região.

pf: Preço da dotação de fator nacional.

Fonte: Adaptado de Gurgel et al. (2013).

Figura 2: Árvore tecnológica da estrutura do bloco de alocação de fatores do PAEG.

Contudo, esse bloco permite representar algum grau de mobilidade de fatores entre regiões, na medida em que uma mudança na remuneração relativa de uma região em relação às demais tende a atrair trabalho e capital das demais regiões do país. Também é possível alterar as elasticidades de transformação a fim de representar a hipótese de livre mobilidade dos fatores entre as regiões brasileiras ($\sigma = \infty$), de tal forma que, após um choque, qualquer diferença na remuneração de um fator entre regiões é completamente eliminada pela migração de fatores, o que significa que há apenas um preço (salário ou retorno do capital) em todas as regiões do país. Para representar a mobilidade perfeita de fatores, a elasticidade de transformação σ , é equivalente a ∞ , uma vez que se especifica no modelo, em substituição à árvore tecnológica representada na Figura 2, funções de transformação da dotação de fator de cada região, precificada por $pfbra(f,r)$, em um fator de produção de uso nacional, precificado como $pf(f)$, sem a atribuição regional ao mesmo. Isso significa que, após um choque, o total utilizado de capital e trabalho em uma dada região não precisa ser igual à dotação inicial destes fatores, mantendo-se, contudo a consistência agregada a nível nacional, de que a soma do fator utilizado nas cinco regiões brasileiras seja igual à soma da dotação inicial do fator nas regiões. Por fim, pode-se também considerar a hipótese de não mobilidade dos fatores produtivos entre as regiões brasileiras, o que significa que cada fator de produção é específico da região, podendo apenas ser transferido ou alocado entre setores dentro da própria região.

O problema de otimização na produção de Y_{ir} define uma função de produção com elasticidade de substituição constante (CES), em que componentes do valor adicionado (fatores primários de produção) podem ser substituídos, sendo tal processo determinado a partir de uma elasticidade de substituição representada pelo parâmetro $esubva_j$ no modelo. Os insumos intermediários e o valor adicionado são combinados a partir de uma função Leontief, em que não podem ser substituídos uns pelos outros. Cada insumo intermediário j , nessa função Leontief, é uma combinação entre uma parcela doméstica e importada do mesmo bem j , a partir de uma função CES de elasticidade de substituição, representada pelo parâmetro $esubd_i$.

O consumo da administração pública é representado, no modelo, por uma agregação Leontief, composta por bens domésticos e importados. O consumo do agente privado pode ser representado por um problema de minimização do custo de dado nível de consumo agregado.

No modelo PAEG, a mensuração dos resultados é dada através de parâmetros e de cálculos dos impactos do cenário implementado. Variação Equivalente (Ev) é o nome que recebe o parâmetro que armazena o resultado da mudança percentual no bem-estar.

O fechamento do modelo PAEG considera fixa a oferta total de cada fator de produção, mas garante mobilidade entre os setores, dentro de uma região. A mobilidade dos fatores primários entre as regiões brasileira pode ser total, parcial ou inexistente, e o presente estudo fará uma análise para as três situações. O modelo considera que não há desemprego; portanto, os preços dos fatores são flexíveis. Pela ótica da demanda, investimentos e fluxos de capitais são mantidos fixos, bem como o saldo do balanço de pagamentos. Dessa forma, mudanças na taxa real de câmbio devem ocorrer para acomodar alterações nos fluxos de exportações e importações após os choques. O consumo do governo poderá alterar com mudanças nos preços dos bens, assim como a receita advinda dos impostos estará sujeita a mudanças no nível de atividade e no consumo. A Tabela 2 mostra como os setores e as regiões do modelo foram agregados nesse estudo.

Tabela 2: Agregação dos setores do modelo PAEG.

Setores	Siglas	Regiões	Sigla
Arroz	(pdr)	Brasil – Região Norte	NOR
Milho e outros cereais	(gro)	Brasil – Região Nordeste	NDE
Soja e outras oleaginosas	(osd)	Brasil – Região Centro-Oeste	COE
Cana-de-açúcar, beterraba, indústria do açúcar	(c_b)	Brasil – Região Sudeste	SDE
Carnes e animais vivos	(oap)	Brasil – Região Sul	SUL
Leite e derivados	(rmk)	Resto do Mercosul	RMS
Outros produtos agropecuários	(agr)	Venezuela	VEM
Produtos alimentares	(foo)	Estados Unidos	USA
Indústria Têxtil	(tex)	Resto do Nafta	RNF
Vestuário e calçados	(wap)	Resto da América	ROA
Madeira e mobiliário	(lum)	Europa	EUR
Papel, celulose e indústria gráfica	(ppp)	China	CHN
Químicos, indústria da borracha e plásticos	(crp)	Resto do mundo	ROW
Manufaturados	(man)		
Eletricidade, gás, distribuição água	(siu)		
Construção	(cns)		
Comércio	(trd)		
Transporte	(otp)		
Serviços e Administração Pública	(adm)		

Fonte: Resultado da Pesquisa.

2.2. Cenário Analítico

O cenário analítico estruturado considera mobilidade total, parcial e ausência de mobilidade dos fatores produtivos (trabalho e capital) entre as regiões brasileiras. Essa alteração na mobilidade de fatores é obtida modificando-se a elasticidade de transformação, σ , do bloco de produção *ftr* do modelo. Ao se considerar $\sigma = 0$, admite-se uma função de transformação do tipo Leontief, que representa a ausência de mobilidade entre os fatores produtivos, ou seja, eles são combinados em proporções fixas,

antes e após o choque do modelo; Quando $\sigma = 1$, considera-se uma função de transformação do tipo Cobb-Douglas, que representa a combinação de fatores que compõe a mobilidade parcial. No último caso o parâmetro σ é deixado livre ($\sigma = \infty$), representando a mobilidade total dos fatores.

A fim de mensurar o efeito da política de ETJ sobre as macrorregiões brasileiras, eliminam-se todo o gasto governamental com a política de ETJ, isto é, com o subsídio da ETJ, e todo o crédito rural subsidiado disponibilizado pela política de ETJ. A partir de dados referentes ao gasto do governo com a ETJ para cada produto agropecuário em cada macrorregião, calcula-se uma proporção dos subsídios totais que compete à ETJ para cada cultura e região. É promovido um choque na variável subsídio (*rto*) apenas nas atividades do setor agropecuário. Esse choque simula a completa eliminação da ETJ. Em algumas atividades de determinadas regiões, as alíquotas encontradas para a ETJ são maiores que a alíquota do subsídio total representada no PAEG, sendo assim, quando a alíquota da ETJ supera a alíquota referente ao subsídio total no PAEG, considera-se que todo subsídio naquele setor consiste em ETJ e, portanto, todo ele é eliminado; entretanto, quando a alíquota de ETJ não supera a alíquota de subsídio total em determinada atividade no PAEG, retira-se apenas a parcela referente a ETJ do subsídio total.

Após a implementação do choque que retira o subsídio da ETJ do setor agrícola, o crédito disponibilizado pelo subsídio da ETJ é excluído da agropecuária e permite-se que seja realocado livremente entre os setores (incluindo os agropecuários), de acordo com a atratividade dos mesmos.

A fim de que o crédito disponibilizado pelos bancos por meio da ETJ seja retirado da agropecuária e realocado na economia, de acordo com a atratividade dos setores, cria-se um artifício na modelagem. Esse artifício consiste em adicionar um novo fator de produção fixo (artificial) aos setores receptores da ETJ, na proporção de apenas 1% do total do valor da produção de cada setor receptor do crédito subsidiado pela ETJ, para não distorcer a contabilidade do setor, mas que deve ser considerado como complementar perfeito (Leontief) ao agregado de demais insumos e fatores de produção utilizados pelo setor. Então, quando do choque de retirada do crédito, diminui-se a oferta desse fator de produção artificial na mesma proporção em que o setor recebe o crédito da ETJ. Como exemplo, se o crédito disponibilizado pela ETJ para um setor qualquer na região Sudeste for calculado como 10% do valor da produção do setor no ano base do modelo, diminui-se a oferta do fator de produção fixo artificial daquele setor em 10%. Como tal fator possui uma relação complementar perfeita com os demais insumos e fatores utilizados pelo setor, isso garantirá que o setor reduza todo o volume de recursos utilizados para sua produção em 10%. Esses recursos podem, então, ser utilizados em qualquer outro setor da economia.

Para garantir que os próprios setores agropecuários que recebem ETJ também possam ser receptores desse volume de crédito, caso sejam suficientemente competitivos, permite-se que o fator fixo artificial seja produzido a partir de uma função que combina capital e trabalho. As proporções de capital e trabalho nesse setor são as mesmas desses fatores no estoque total de fatores da região. Dessa forma, se um determinado setor agropecuário ainda for relativamente mais competitivo e atrativo que outros, mesmo sem o subsídio da ETJ, a eliminação do subsídio e a retirada forçada do crédito associado ao subsídio àquele setor não impedirão que o setor volte a crescer, já que o fator fixo artificial pode ser gerado pela combinação de capital e trabalho.

3. RESULTADOS

Nessa seção é apresentada a análise de resultados para a simulação do cenário proposto nesse estudo.

3.1 Efeitos dos gastos do governo e do crédito rural disponibilizado pela ETJ com realocação desse crédito nas economias regionais

Nessa seção objetiva-se implementar um choque em que retira-se dos setores agrícolas o subsídio da ETJ e o crédito disponibilizado via ETJ. Esse crédito disponibilizado pela ETJ será realocado livremente entre os setores (incluindo o agropecuário), de acordo com a atratividade de cada um deles. Todos os resultados serão apresentados e discutidos com os sinais trocados, representando o efeito que teria a introdução do subsídio via ETJ e do volume de crédito rural disponibilizado por indução da aplicação da ETJ exclusivamente para os setores agropecuários. Esses resultados representam o retorno do subsídio e do crédito rural no seu melhor emprego alternativo. A análise é feita considerando mobilidade nula, parcial e total dos fatores de produção entre as regiões brasileiras.

3.1.1 Impactos no PIB e no retorno aos fatores primários

A fim de avaliar o retorno da política são examinados os efeitos do subsídio e do crédito rural disponibilizado pela ETJ em termos de sua capacidade de promover crescimento econômico e bem-estar nas regiões brasileiras. A Tabela 3 mostra os resultados para variações no PIB das regiões brasileiras, em termos monetários, comparado aos gastos do governo com a política de ETJ.

Verifica-se que quando há ausência de mobilidade entre os fatores a ETJ gera um aumento no PIB das regiões brasileiras, exceto para o Sudeste. A região Nordeste foi a maior beneficiada com a política. O gasto nessa região é de R\$ 0,33 bilhão, disponibilizando R\$ 1,51 bilhão de crédito rural. A ETJ, e o crédito subsidiado, geraram um aumento no PIB de R\$ 0,13 bilhão. O efeito multiplicador nessa região mostra que para cada real gasto com a política ETJ, há um aumento de R\$ 0,40 no PIB.

Tabela 3: Efeitos dos gastos com equalização das taxas de juros dos recursos aplicados sob a forma de crédito rural no PIB das regiões brasileiras, na ausência de mobilidade de fatores, 2007 (R\$ bilhões).

Regiões	Gasto ETJ ¹	Crédito Gerado	Mobilidade Nula		Mobilidade Parcial		Mobilidade Total	
			Efeito no PIB ²	Multipli-Cador ^{2/1}	Efeito no PIB ²	Multipli-Cador ^{2/1}	Efeito no PIB ²	Multipli-Cador ^{2/1}
NOR	0,11	0,50	0,02	0,18	-0,26	-2,45	-0,69	-6,41
NDE	0,33	1,51	0,13	0,40	-0,25	-0,78	-0,3	-0,92
COE	0,42	2,07	0,08	0,20	0,01	0,03	0,5	1,21
SDE	0,97	4,79	-0,08	-0,08	0,47	0,49	-0,59	-0,61
SUL	1,07	5,13	0,11	0,10	0,36	0,34	1,47	1,37
BRA	2,89	14,02	0,26	0,09	0,33	0,11	0,38	0,13

Fonte: Resultados da pesquisa.

Com mobilidade parcial, os resultados mostram que o recurso destinado à política ETJ e o crédito subsidiado geram aumento no PIB de três das cinco regiões analisadas. A região Sudeste foi a que apresentou o resultado mais positivo, tanto em termos de PIB, R\$ 0,47 bilhão, quanto de multiplicador. Para cada real gasto com ETJ, há um aumento no PIB de R\$ 0,49.

A total mobilidade dos fatores produtivos entre as regiões gera diferentes efeitos no PIB em relação à análise feita com ausência de mobilidade. Quando há remoção do gasto com ETJ, apenas as regiões Centro Oeste e Sul respondem positivamente. Enquanto que para as regiões Norte, Nordeste e Sudeste o efeito sobre o PIB é negativo. Esses resultados mostram-se diferentes do trabalho apresentado por Cardoso *et al* (2014) apenas para a região Nordeste, onde os autores, em um estudo feito para o ano de 2004, com total mobilidade de fatores encontram uma variação positiva de R\$ 0,97 bilhão para o PIB da região Nordeste.

Na ausência de mobilidade entre os fatores de produção, o gasto com ETJ, de modo geral, gera crescimento no PIB menor que o custo da política, com aumento de R\$ 0,26 bilhão para o Brasil. A presença da política ETJ, em termos de geração de crescimento econômico, apresenta taxa de retorno negativa, ou seja, para cada real gasto com ETJ no Brasil, há um aumento de R\$ 0,09 na economia.

Quando há mobilidade parcial dos fatores produtivos, o resultado agregado, para o Brasil, mostra que a política gera crescimento econômico em termos de PIB, R\$ 0,33 bilhão, com taxa de retorno negativa, ou seja, para cada real gasto com a política ETJ, há um aumento de R\$ 0,11 no PIB brasileiro.

Considerando a total mobilidade dos fatores de produção, percebe-se que no agregado Brasil, a política gera efeitos positivos sobre o PIB. Mas não se afirma o mesmo para todas as regiões brasileiras. Sendo assim, sob esse efeito de mobilidade dos fatores, não se pode afirmar que a política é favorável à economia de todas as regiões.

O gasto com ETJ e o crédito subsidiado mostram-se importantes para o PIB das regiões Sul e Centro Oeste, sob os diferentes movimentos de fatores adotados. Esse resultado indica o quanto o setor agrícola é forte nessas duas regiões e fortalece a ideia de que o subsídio gera crescimento econômico.

Para as regiões Norte e Nordeste, na análise onde não há mobilidade entre os fatores de produção o subsídio se mostra importante, gerando crescimento econômico para ambas as regiões. Quando há mobilidade parcial e total de fatores, a ETJ e crédito subsidiado geram retração econômica em ambas as regiões, o que indica que nessa situação, há distorções da aplicação do subsídio. Conclui-se que o estímulo gerado pelo subsídio da ETJ em termos de competitividade dos setores agropecuários do Sul e do Centro-Oeste devem tornar essas regiões bem mais atrativas para a produção agrícola. Assim Norte e Nordeste acabam perdendo capital e trabalho para o Sul e o Sudeste, impactando em queda na produção, perda de arrecadação do governo e queda nas exportações da região.

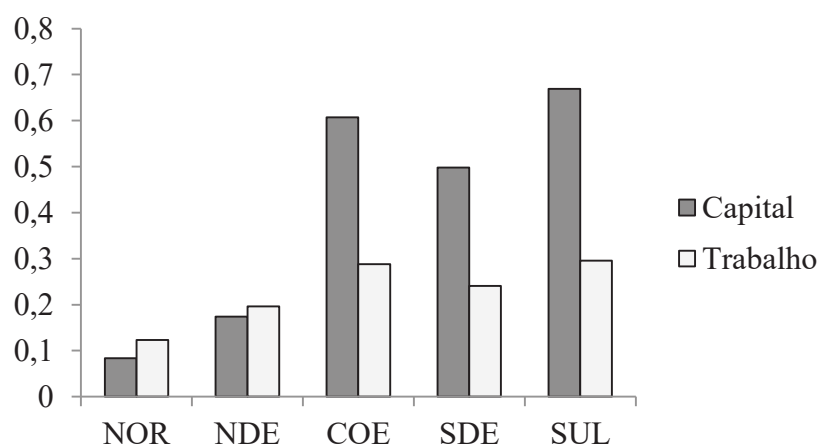
A região Sudeste não mostrou crescimento econômico frente aos gastos com ETJ e crédito subsidiado, tanto com ausência quanto com total mobilidade entre os fatores de produção. Esse é um indicativo de que nessa região, sob essas condições de mobilidade, poderia haver aplicações mais eficientes dos recursos dispendidos com ETJ ao setor agrícola e fortalece a ideia da necessidade de análises regionais mais profundas, afim de que se possam desenvolver políticas econômicas que consigam suprir as especificidades de cada região. Entretanto, ao se analisar a mobilidade parcial, a região foi a maior beneficiada pela política.

Os resultados encontrados contrariam a hipótese inicial dessa pesquisa e podem ser explicados pelo padrão relativo de competitividade das diferentes regiões brasileiras, mediante as diferentes análises, de mobilidade de capital e trabalho adotadas.

Considerando os diferentes efeitos para as mobilidades dos fatores de produção analisadas nesse estudo, e lembrando que o movimento dos fatores é uma alternativa ao comércio de bens e serviços, as Figura 3 e 4 mostram a mudança em termos monetários, do retorno ao capital e ao salário, pagos para cada região mediante a política ETJ e crédito subsidiado, com ausência e total mobilidade respectivamente.

A Figuras 3 mostra que o gasto com ETJ e o crédito subsidiado, geram aumentos no retorno ao capital e ao salário pagos, em relação ao *benchmark*. Os maiores resultados são das regiões Sul e Centro Oeste, seguidas pelo Sudeste. Como não há mobilidade regional entre os fatores de produção, as regiões que recebem maiores subsídios ao setor agrícola, e com isso maior crédito subsidiado, vão perceber maiores efeitos.

Considera-se que a ausência de mobilidade entre os fatores produtivos é um limitador que pode subestimar os resultados, uma vez que a realocação do volume de crédito que circula na economia, antes subsidiado, não consegue ser absorvido de maneira eficiente pelos setores nas regiões brasileiras.



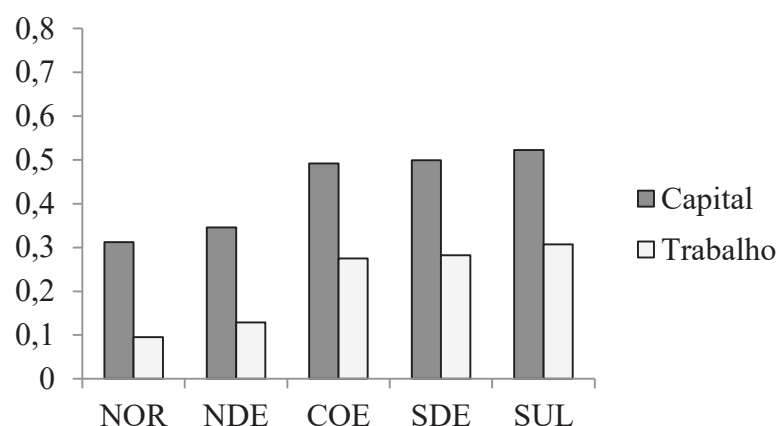
Fonte: Resultados da pesquisa.

Figura 3: Efeito dos gastos com ETJ e crédito subsidiado sobre a variação do retorno ao capital e massa salarial pagos nas regiões brasileiras considerando ausência de mobilidade dos fatores produtivos, 2007 (%).

Pela Figura 4, a política ETJ e o crédito subsidiado ao setor agrícola, em um cenário onde os fatores produtivos circulam livremente entre as regiões brasileiras, geram expansão no retorno ao capital e ao salário pagos, em relação ao *benchmark*.

Essas mudanças nas remunerações dos fatores são oriundas das forças de demanda por fatores, dada pelos setores econômicos demandando mais fatores produtivos, e de oferta de fatores, dada pela dotação inicial e possibilidade de mobilidade entre regiões. Uma maior valorização dos fatores em termos reais sugere um maior aquecimento das atividades econômicas da região e, possivelmente, maior atração de recursos.

A total mobilidade dos fatores de produção entre as regiões brasileiras atenua os efeitos de expansão do retorno ao capital e ao salário, que são mais acentuados quando há ausência de mobilidade, já que após o choque, capital e trabalho migrarão para regiões onde há melhor remuneração. Com o aumento da oferta de fatores, frente à demanda dos setores de cada região, haverá queda da remuneração até que se atinja o equilíbrio.



Fonte: Resultado da pesquisa.

Figura 4: Efeito dos gastos com ETJ e crédito subsidiado sobre a variação do retorno ao capital e salários pagos nas regiões brasileiras, com total mobilidade dos fatores produtivos, 2007 (%).

Esses resultados estão em consonância com os de Cardoso *et al* (2014) para as regiões Centro-Oeste, Nordeste e Sul, onde os autores encontram acréscimos no retorno ao capital e ao salário pagos sob a política de ETJ. Os autores encontram retração para as regiões Norte e Sudeste, sob a justificativa de que como os subsídios distorcem a produção setorial, as regiões em que o setor de manufaturados detém participação superior à agricultura na economia, percebem desestímulo (relativo) causado pelo estímulo (relativo) ao setor agrícola, e este vai provocar mobilidade dos fatores produtivos para as regiões em que a agricultura representa parcela maior (Centro-Oeste) ou próxima à participação da indústria na atividade econômica (Nordeste e Sul). Essas regiões vão absorver os fatores migrantes, principalmente no setor agrícola, que é a atividade estimulada pelo incentivo, e, assim, aumentarão seu nível de atividade.

A análise do retorno ao Capital com mobilidade parcial de fatores está representada na Tabela 4, que mostra as variações em termos percentuais no Índice de Preços ao Consumidor (IPC), na oferta de capital empregado e retorno ao capital pago nas regiões brasileiras. O resultado mostra que se o retorno do capital está aumentando, a rentabilidade do capital cresceu mais, ou decresceu menos, que o IPC. Observa-se que há aumento no IPC de todas as regiões.

Tabela 4: Efeitos da política ETJ e dos recursos aplicados sob a forma de crédito rural sobre a variação do IPC, da oferta e da remuneração do capital das regiões brasileiras, 2007 (%).

Regiões	IPC	Oferta de Capital	Capital Fora *	Capital Local**
NOR	0,124	-0,305	0,161	0,003
NDE	0,156	-0,193	0,236	0,099
COE	0,116	0,051	0,512	0,055
SDE	0,134	0,053	0,511	0,790
SUL	0,133	0,133	0,581	0,297

Fonte: Resultados da pesquisa.

*Refere-se ao preço (ou remuneração) do fator que é pago pelos setores que usam o fator. O preço de equilíbrio é dado pelas forças de oferta do fator, que incluem o fator total disponível na região, incluindo a parcela que migrou para a mesma, e pelas forças de demanda do fator, advindas dos setores produtivos.

**Refere-se ao preço (ou remuneração) do fator que é recebido pelas famílias originais da região. Esse preço é consequência do equilíbrio entre o total de fator originalmente ofertado pela região (estoque inicial do fator) e a demanda originada na função de transformação dos fatores.

Nas regiões Norte e Nordeste há uma queda da oferta de capital, mostrando que há menos atratividade para o fator capital nessas regiões em relação às outras. Essa queda da oferta de capital, frente à demanda dos setores por esse fator, gera um aumento na remuneração do fator. Como esse aumento foi superior ao aumento no IPC, as famílias detentoras de capital que já estavam nessas regiões, perceberam um pequeno aumento real na remuneração do capital.

Nas regiões Centro Oeste, Sudeste e Sul há aumento na oferta de capital. Apesar do aumento na oferta do fator nessas regiões, há uma variação positiva em sua remuneração. Esse é um indicativo, que Centro Oeste, Sul e Sudeste, que são as regiões que recebem maior volume de crédito subsidiado, são mais sensíveis ao choque da ETJ. As famílias detentoras de capital, que já se encontravam nessas regiões, percebem variações positivas na remuneração do capital.

A Tabela 5 mostra as variações em termos percentuais no IPC, na oferta de trabalho e retorno do salário pago nas regiões brasileiras.

Tabela 5: Efeitos da política ETJ e dos recursos aplicados sob a forma de crédito rural sobre a variação do IPC, da oferta de trabalho e do salário pago nas regiões brasileiras, 2007 (%).

Regiões	IPC	Oferta de Trabalho	Trabalho Fora*	Trabalho Local**
NOR	0,124	-0,174	0,088	-0,086
NDE	0,156	-0,068	0,156	-0,032
COE	0,116	-0,014	0,243	-0,032
SDE	0,134	0,032	0,287	0,455
SUL	0,133	0,048	0,293	0,118

Fonte: Resultados da pesquisa.

*Refere-se ao preço (ou remuneração) do fator que é pago pelos setores que usam o fator. O preço de equilíbrio é dado pelas forças de oferta do fator, que incluem o fator total disponível na região, incluindo a parcela que migrou para a mesma, e pelas forças de demanda do fator, advindas dos setores produtivos.

**Refere-se ao preço (ou remuneração) do fator que é recebido pelas famílias originais da região. Esse preço é consequência do equilíbrio entre o total de fator originalmente ofertado pela região (estoque inicial do fator) e a demanda originada na função de transformação dos fatores.

Observa-se uma queda na oferta de trabalho nas regiões Norte, Nordeste e Centro Oeste. Essa queda, frente à demanda dos setores por este fator, provoca aumento no salário pago pelos setores produtivos nessas regiões. As famílias da região, por sua vez, experimentam uma pequena queda na renda do trabalho, portanto a mão-de-obra local percebe uma desvalorização de sua remuneração em relação ao valor da cesta de consumo da região. O resultado para a região Centro-Oeste reflete a característica do setor agropecuário da região ser mais intensivo no uso do fator capital.

Sudeste e Sul apresentam expansão em sua oferta de trabalho. Essas regiões se mostram mais sensíveis ao choque da política ETJ e crédito subsidiado, apresentando aumento no salário pago. As famílias da região experimentam um aumento na renda do

trabalho, com isso a mão-de-obra local percebe uma valorização de sua remuneração em relação ao valor da cesta de consumo da região.

Quando se assume a mobilidade parcial de fatores, Capital e Trabalho se movem entre as regiões, porém de maneira limitada, diferente de uma situação onde há total mobilidade dos fatores. A mobilidade total entre as regiões permite que os fatores se movam até equalizar suas respectivas remunerações entre as regiões. Portanto, nas regiões em que o fator se tornaria mais escasso, e com isso mais demandado, há maior atração desse fator vindo de outras regiões, até que o aumento da sua oferta na região mais atrativa e a redução nas regiões menos atrativas, permita equilibrar sua remuneração em um mesmo nível para todas as regiões. Para a mobilidade parcial, existem restrições econômicas e institucionais que impedem a completa equalização da remuneração desses fatores entre as regiões. Sendo assim, a remuneração recebida por um mesmo fator difere entre as regiões, o que significa que o fator que permaneceu em uma dada região pode receber uma remuneração diferente do fator que saiu desta região e migrou para outra.

3.1.2. Impactos dos gastos com ETJ sobre o bem-estar das regiões

A política de subsídios tem impacto direto sobre o consumo dos agentes e com isso, sobre seu bem estar⁴, já que influencia a quantidade de bens produzida, o fluxo de exportações e importações, os preços domésticos e de importados, e a renda da economia. A Tabela 6 mostra as variações de bem estar para cada região brasileira.

Os resultados mostram que há aumento de bem estar na presença da política ETJ, já que com o subsídio haverá queda nos preços dos produtos agrícolas e, com isso, aumento no consumo dos mesmos, além de aumento no Fluxo Comercial e no consumo intermediário.

Todas as regiões do modelo apresentam crescimento em termos de bem estar. Para o Brasil, a análise com as três mobilidades propostas apresentou variações positivas de bem estar, e muito próximas entre si. Na ausência de mobilidade a variação foi de R\$ 6,12 bilhões, com total mobilidade R\$ 6,19 bilhões e com mobilidade parcial R\$ 6,15 bilhões. Sendo assim, a política melhorou a situação do Brasil em termos de consumo, o mesmo ocorreu para todas as regiões brasileiras.

Tabela 6: Efeitos dos gastos com ETJ e recursos aplicados sob a forma de crédito rural, medidos pela variação equivalente, sobre o bem-estar das regiões brasileiras, 2007 (Variação Equivalente em R\$ bilhões).

Regiões	Gasto com ETJ ¹	Crédito Gerado	Mobilidade Nula		Mobilidade Parcial		Mobilidade Total	
			R\$ Bilhões ²	Multiplificador ^{2/1}	R\$ Bilhões ³	Multiplificador ^{3/1}	R\$ Bilhões ⁴	Multiplificador ^{4/1}
NOR	0,11	0,50	0,09	0,83	-0,02	-0,19	0,21	1,95
NDE	0,33	1,51	0,38	1,16	0,10	0,30	0,53	1,61
COE	0,42	2,07	0,67	1,61	0,03	0,07	0,56	1,34
SDE	0,97	4,79	3,13	3,22	5,24	5,39	3,29	3,38
SUL	1,07	5,13	1,86	1,74	0,81	0,76	1,59	1,49
BRASIL	2,89	14,02	6,12	2,11	6,15	2,13	6,19	2,14

⁴ Ocorre uma alteração no nível de bem estar dos agentes quando há alterações em seu consumo de bens e serviços. No modelo, há aumento de bem estar de um agente quando esse aumenta seu nível de consumo, uma vez que representa-se a função de utilidade considerada no estudo é considerada como uma função Cobb-Douglas do consumo dos bens constantes na cesta de demanda agregada das famílias de cada região.

Os resultados do efeito multiplicador do bem estar sobre as regiões brasileiras e sobre o Brasil é negativo apenas para a região Norte quando se considera mobilidade parcial dos fatores produtivos. Para todas as outras análises, o multiplicador do bem-estar é positivo.

Esses resultados vão de acordo com os encontrados por Cardoso *et al* (2014), os autores mostram que o gasto com a política de ETJ traz ganhos em bem-estar para todas as regiões, tendo em vista que o consumo aumenta para todas elas, mesmo no Norte e no Sudeste, cujo efeito total sobre o PIB encontrado por eles é negativo. O bem-estar é maior na presença da ETJ porque subsídios a produtos agrícolas incentivam o aumento na produção e na elevação do consumo pela redução no preço.

Os resultados mostram que quando se considera a ausência de mobilidade entre os fatores produtivos, com exceção da região Norte, todas as demais regiões apresentam um efeito multiplicador que gera retornos positivos, ou seja, para cada real gasto com a política, há um retorno maior que R\$ 1,00 sobre o bem-estar.

O bem-estar é determinado pelo consumo das famílias, portanto alimentos são um importante componente da cesta de consumo, principalmente nas regiões Norte, Nordeste e Centro-Oeste. Dessa forma, apesar do subsídio implicar em distorção, ele permite um fluxo de recursos subsidiados ao setor agropecuário, bem como um fluxo de recursos extras, que aumentam a oferta dos bens agropecuários e reduzem os preços no setor de alimentos, implicando em ganhos para o consumo e bem-estar das famílias.

4. CONCLUSÃO

O presente estudo busca contribuir com o debate relacionado à intervenção estatal na economia, analisando como a mobilidade dos fatores primários, capital e trabalho, afeta o crescimento econômico e o bem-estar gerados pela política de equalização das taxas de juros (ETJ).

Para desenvolver o estudo utilizou-se o conjunto metodológico do Projeto de Análise de Equilíbrio Geral da Economia Brasileira (PAEG). Foi analisado um cenário onde removem-se do ambiente econômico os gastos com a ETJ e todo o crédito gerado pela política ao setor agrícola e posteriormente realoca-se esse crédito entre todos os setores da economia (inclusive o agrícola) de acordo com a atratividade de cada um. A análise é feita com ausência de mobilidade, mobilidade parcial e total dos fatores produtivos.

Ao se analisar a mobilidade dos fatores produtivos, conclui-se que, em termos de PIB, a região Sul e o Brasil, respondem positivamente ao aumento da mobilidade, ou seja, quanto maior a mobilidade dos fatores, maior a variação do PIB dessa região e do Brasil. O contrário acontece com as regiões Norte e Nordeste, ou seja, quanto maior a mobilidade entre os fatores, maiores são as variações negativas sobre o PIB dessas regiões. A região Centro Oeste responde mais positivamente quando há mobilidade total entre os fatores produtivos e a região Sudeste quando há mobilidade parcial dos fatores. Essas duas regiões não apresentaram um comportamento que caminhe de acordo com o aumento da mobilidade.

A política de ETJ promove variação positiva do PIB para a maioria das regiões, e para o Brasil como um todo, mas apresenta taxa de retorno negativa em termos de geração de crescimento econômico. Todas as regiões respondem positivamente, em termos de bem-estar, ao efeito da política de ETJ. Assim como na análise do PIB, a variação de bem-estar das regiões não apresenta um padrão regular de acordo com a mobilidade dos fatores. A política ETJ gera aumento de bem estar maior que seu custo para o Brasil.

Conclui-se que algumas políticas governamentais podem se mostrar eficientes em termos de gerar crescimento do PIB e do bem-estar. A pesquisa, portanto, cumpriu com o objetivo de analisar como a mobilidade dos fatores produtivos afeta o crescimento econômico e o bem-estar gerados pela política de equalização de taxa de juros nas regiões brasileiras.

REFERÊNCIAS BIBLIOGRÁFICAS

- CARDOSO, D. F.; TEIXEIRA, E. C.; GURGEL, A. C.; de CASTRO, E. R. Effects of Government Intervention Via Rural Credit Subsidy on Economic Growth and Welfare of Brazilian Regions. *Indian Journal of Agricultural Marketing*, v. 25, p. 41-50, 2011.
- CARDOSO, D. F.; TEIXEIRA, E. C.; GURGEL, A. C.; de CASTRO, E. R. Intervenção governamental, crescimento e bem-estar: efeitos da política de Equalização das Taxas de Juros do crédito rural nas regiões brasileiras. *Nova Economia (UFMG. Impresso)*, v. 24, p. 363-388, 2014.
- CASTRO, E. R.; TEIXEIRA, E. C. Retorno dos gastos com a equalização das taxas de juros do crédito rural na economia brasileira. *Revista de Política Agrícola*. Ano 3, n. 3, Jul./Ago./Set. 2004. p. 52 a 57.
- CODSI, G., PEARSON, K. R. GEMPACK: general-purpose software for applied general equilibrium and other economic modellers. *Computer Science in Economics and Management*. v.1, p.189-207, 1988.
- GASQUES, J. G.; VILLA VERDE, C. M. Gastos Públicos na Agricultura: Evolução e Mudanças. Texto para discussão n. 948. Brasília: IPEA, abr. 2003. 31 p.
- GLOBAL TRADE ANALYSIS PROJECT – GTAP, 2001. (<http://www.agecon.purdue.edu/gtap/>). Hertel, T. W. (ed.) *Global trade analysis: modeling and applications*. Cambridge University Press, Cambridge and New York, 1997.
- GURGEL, A.C.; PEREIRA, M.W.G.; TEIXEIRA, E.C. A estrutura do PAEG. PAEG. Technical Paper No.1 e No.5. Viçosa: DER/UFV. (2011).
- HERTEL, T. W. (ed.) *Global trade analysis: modeling and applications*. Cambridge University Press, Cambridge and New York, 1997.
- MUNDELL, R. A. A Theory of Optimum Currency Areas. *The American Economic Review*, v. 51, n. 4, p. 657-665, set. 1961.
- RUTHERFORD, T. F., PALTSEV, S. V. *GTAPinGAMS and GTAP-EG*: global datasets for economic research and illustrative models. Working Paper, Department of Economics, University of Colorado, 64 p., 2000. (<http://nash.colorado.edu/gtap/gtapgams.html>).
- RUTHERFORD, T. F. *GTAP6inGAMS*: The dataset and static model. 42 p., 2005, mimeo. (<http://www.mpsge.org/gtap6/gtap6gams.pdf>).
- SOUZA, N. de J. Economia Regional: Conceitos e Fundamentos Teóricos. *Perspectiva Econômica*. XVI, v.11, n. 32, 1981, p. 67-102.
- TAYLOR, L. *Income distribution, inflation and growth*. Cambridge, Massachusetts: The MIT Press, 1994. p. 183-196.
- TEIXEIRA, E. V.; PEREIRA, M. W. G.; GURGEL, A. C. *A Estrutura do PAEG*. 1ª ed. Campo Grande, 2013. 198 p.

THE INNOVATIVE ASPECTS OF BITCOIN, MARKET MICROSTRUCTURE AND RETURNS VOLATILITY: AN APPROACH USING MGARCH

Bruno Ferreira Frascaroli¹
Thiago Carvalho Pinto²

ABSTRACT

This study seeks to analyze conceptual, innovative, marketing and quantitative aspects of Bitcoin (BTC) and how these are reflected in the volatility of its return. After describing basic concepts of digital currencies and BTC, an electronic currency created in 2009, we contextualize BTC as a financial innovation. Regarding market-oriented and structural aspects, we investigated the existence of similarities between BTC's credit and debit market and the traditional structure of the payment card market, known as the Two-Sided Market (2SM). The BTC's perfect adequacy for the 2SM structure was established, but only with respect to the debit-card market. BTC returns for the period from September 2011 to June 2015 were then used as a sample, and returns for the period from March 2013 to June 2015 were used as a subsample. Based on these data, the DCC MGARCH model was estimated, making it possible to obtain the parameters and quasi-correlations of volatilities among variables. The Standard and Poor's 500 index (SPX-500), the price of gold and China's main stock index (SSEC) were considered by virtue of previous exercises conducted by Pavel, D'Artis and Miroslava (2014) and Stråle and Tjernström (2014). The results showed the existence of persistent volatility, possibly indicating that BTC returns had two different phases: (1) a period of “euphoria” followed by (2) a period of “convergence” to the fundamentals that determine these returns.

Keywords: Bitcoin. Innovation. 2SM. Volatility. DCC MGARCH.

RESUMO

A presente pesquisa buscou analisar aspectos conceituais, inovativos, mercadológicos e quantitativos do Bitcoin (BTC) e como os mesmos se traduziram na volatilidade dos seus retornos. Após expostos conceitos basilares sobre as moedas digitais e o BTC, moeda eletrônica criada em 2009, buscou-se contextualizá-la como inovação financeira. Quanto aos aspectos mercadológicos e estruturais, foi investigada a existência de semelhanças entre o mercado de crédito e débito do Bitcoin e a estrutura tradicional do mercado de cartões, conhecida por Mercado de Dois Lados (M2L). Verificou-se a perfeita adequação do BTC à estrutura M2L, mas apenas no mercado de cartões de débito. Mais adiante, foram usadas como amostra os retornos do BTC do período de Setembro de 2011 a Junho de 2015 e, como subamostra, a janela temporal a partir de Março de 2013 a Junho de 2015. A partir desse conjunto de informações, estimou-se o modelo DCC MGARCH, que possibilitou a obtenção dos parâmetros e das quasicorrelações das volatilidades entre as variáveis. Foram consideradas o índice *Standard and Poor's 500* (SPX-500), a cotação do ouro e o principal índice da bolsa da China (SSEC), em virtude de exercícios anteriores realizados por Pavel, D'Artis e Miroslava (2014) e Stråle e Tjernström (2014). Os resultados apontaram existência de volatilidade persistente e que mostra, possivelmente, que os retornos da moeda tiveram duas fases: primeiramente um momento de ‘euforia’ e depois de ‘convergência’ aos fundamentos que determinam os mesmos.

Palavras-Chave: Bitcoin. Inovação. M2L. Volatilidade. DCC MGARCH.

JEL Code: E51, G23, C58.

Área 3: Microeconomia Aplicada.

¹ Professor of PPGE, Federal University of Paraíba, Brazil. e-mail: frascaroli.b@gmail.com.

² Master in Economics - PPGE, Federal University of Paraíba, Brazil. E-mail: thiagao.cp@gmail.com.

1 INTRODUCTION

A stable price system is essential to the fulfillment of contracts, to issue securities (private or public) and even to the existence and functioning of the financial market. In this regard, the tasks of regulating the financial system, issuing a currency and maintaining its purchasing power are traditionally assigned to the State. Therefore, monetary policies affect social wellbeing by changing a currency's fiduciary power.

That notwithstanding, globalization has not only enabled the opening of new markets but also strengthened trade ties and created zones of influence. New products and markets are emerging at great speed. More interconnected markets can render the environment more sensitive to fluctuations. It can be argued that financial crises have increased the capacity for contagion; commercial interdependence among countries is worrisome. Accordingly, there are calls for greater regulation of these markets, including the creation of laws and government agencies to instill discipline in the financial system.

Governmental actions can compromise a currency's desirability. Participation in conflicts, import-substitution policies, fixed exchange-rate regimes, economic rescue plans and seigniorage are examples of strategies that involve—at least momentarily, and regardless of success—some degree of currency depreciation to affect the price system.

Mediums of exchange reflect the level of social and technological organization. For example, bartering was the main form of peaceful acquisition of goods and services, but this social dynamic was gradually transformed by socioeconomic, cultural and technological events and processes so that with the passage of time, the use of currencies has dominated other forms of trading and become the norm. Credit and debit cards represent a new technical advancement, providing ease of use, security and the ability to change consumer behavior. The proliferation of digital payment methods (i.e., eWallets) such as Paypal, Google Wallet and MoneyGram has further facilitated e-commerce transactions.

Digital currencies represent an electronic medium of exchange with all of the functions and features of traditional currencies, except that they are not physically tangible. Characteristics such as homogeneity, divisibility, portability and ease of handling and transport are even more evident in digital currencies than in physical currency.

Bitcoin (BTC) is the digital currency with the highest liquidity and daily trading volume on the financial markets. In addition, because it was the first digital currency, there are more observations of its series of quoted prices; moreover, the movements of its series reflect not only quantitative aspects but also social changes. BTC is a non-centralized, encrypted currency³. It is not governed by statute and is not legal tender.

Having presented some of the characteristics of the proposed topic, this study aims to analyze some of the microeconomic and macroeconomic factors that have led to volatility in BTC returns that have two phases of behavior. To this end, BTC is contextualized according to financial innovation concepts and an attempt is made to analyze BTC from a microeconomic perspective using the theoretical structure of two-sided markets (2SM) to identify similarities to and differences from traditional credit and debit systems. Next, based on Pavel, D'Artis and Miroslava (2014) and Stråle and Tjernström (2014), dynamic conditional correlation multivariate generalized autoregressive conditional heteroskedasticity (DCC MGARCH) models were used to estimate the parameters of the volatilities and their quasi-correlations between variables. The Standard and Poor's 500 index (SPX-500), the price of gold and China's main stock index (SSEC) were considered for this purpose.

This article is organized as follows. Following this brief introduction, section 2 presents theoretical concepts related to both innovation and financial innovation. Section 3 presents the characteristics of BTC. Section 4 sets out the theoretical framework of the two-sided market and notes similarities and differences between the payment card market and BTC. Section 5 presents the methods.

³Cryptography is the study of the principles and techniques by which information can be transformed from its original form to another unreadable form so that it can be understood by only the addressee (holder of the "secret key"), thus making it difficult for an unauthorized party to read. Therefore, only the message recipient can easily read the information.

Sample planning and treatment is described in the sixth section. The results obtained are presented in section 7, and section 8 contains the conclusions.

2 INNOVATION ECONOMICS

Innovation economics seeks to identify the causes, dynamics and effects of innovation processes because they transform primary structures, restructuring industries and markets. This section provides a brief review of innovation economics and aims to apply the literature originally conceived for production structures to BTC and other decentralized digital currencies.

2.1 Financial Innovations

According to Jensen et al. apud Oliveira and Torkomian (2009), there are two types of knowledge: *encoded knowledge* and so-called *tacit knowledge*. The first is relatively easy to transmit, whereas the second depends on learning. These concepts are quite common in evolutionary innovation models. Application of these concepts facilitates understanding of the process of market creation in the form of conglomerates.

As a result of the centralization of capital, large financial centers emerge that are geographically clumped together. This conformity depends not only on historical economic and social factors but also on existing externalities in the financial sector. These characteristics promote the maintenance of a production structure in the formation of agglomerates. It can be said that production clusters are knowledge-sharing spaces, emphasizing learning as a localized process that occurs through social interaction and embeddedness, strengthening innovation networks and practices (IPIRANGA, 2006 apud OLIVEIRA; TORKOMIAN 2009, p.4).

The main themes analyzed from the perspective of financial innovation are shared risk and its economic implications, information asymmetry, transaction costs and the role of regulatory restrictions. It is frequently argued that financial markets are excessively affected by intermediaries at the expense of savers and entrepreneurs. The role of intermediaries completes markets by attracting agents and, if economies of scale are found, by lowering costs. Radical financial innovation facilitates the extension of risk management beyond its previous limits, including new risk classes, modifying assumptions regarding financial insurance, hedging strategies and diversification.

The same diffusion process brought about by innovation in other sectors is also observed in financial innovation processes. Schiller (2007, p. 306) argues that financial innovation processes are continuous and tend toward diffusion. The author points to the role of information technology, the understanding of behavior, and psychology applied to finance as determinants of diffusion.

It is understood that full risk sharing, as conceived by Kenneth Arrow, in which there is a market for each price, for each risk and for each natural state, is difficult to measure. Thus, problems such as information asymmetry and its manifestations in the form of *adverse selection* and *moral hazard* associated with risk management involve transaction costs that can be reduced. Wonglimpiyarat (2011) obtains results suggesting that financial innovations affect economic performance by promoting business activities. This is an important sector of the innovation system that can be seen as linking financing and innovation.

The emergence of digital currencies, for example, represents the possibility of disconnecting the financial system from political objectives. It allows users autonomy because they can avoid currencies that have high potential to lose value. Network externalities are also relevant to the increased adoption of user-encrypted currencies because they constitute a special type of externality in which the degree of utility gained by an individual depends on the number of users of the same good. When an individual's willingness to pay exceeds the costs, the market expands and given that at least two equilibria are possible in markets that have network externalities, there is a tendency for the market to reach an equilibrium with a large number of agents.

In markets with network externality characteristics, producers may initially cause disturbances with deals, price discrimination and other strategies to attract users to the network. In the case of digital

currencies, at least three factors may determine the network's ability to be attractive: reliability, the possibility of profit, and the maintenance of an entrant rate until the formation of critical mass, which is given by a of volume users that when achieved, makes the network attractive to non-participants because of that network's very existence.

In addition to BTC, there is an increasing number of other digital currencies. Each such currency emphasizes certain aspects that its creators understand to be the most desirable. Among the most popular include the following:

- *Litecoin* - Although this currency closely resembles BTC, its issuance is constrained to 84 million monetary units, and the difficulty of mining represents something close to a quarter, which allows the use of simpler equipment. The initial appeal of this currency was the relative ease of mining. It is important to note that the mining algorithm of Litecoin is different from that of BTC.
- *Dogecoin* – This currency has unlimited issuance with the clear objective of avoiding becoming a currency with deflationary tendencies and facilitating the eventual formation of a credit market.
- *PPcoin* – This currency offers a nondeterministic money supply and has advantages related to its security module. The return on every mined block depends on mining difficulty, with a new block issued every ten minutes;
- *Dash* (previously known as *Darkcoin*) – This currency aims to make transactions anonymous by increasing tracing difficulty through fictitious transactions.

The initial issuance of BTC at a symbolic value of \$ 0.01 was seen as indicating a profound respect for the market, granting users full power to attribute value through the power of supply and demand. It can also be seen as a promotional pricing strategy because the two understandings are not mutually exclusive. The strategies of initial promotional prices, dissemination strategies and maintenance of a network growth rate are often adopted by companies operating in markets with network externalities, such as the software market.

When considering the initial issuance price and the lower initial difficulty of mining, one can observe strategies that are similar to those adopted by companies that depend on the formation of a network for greater penetration of their products. As a "product" of potential future value, an increasing number of users, attracted by the possibility of gain, form the initial network. Following publication in the media and maintenance of the network's growth rate, the currency appreciates and these movements are fed back until the formation of critical mass, which leads to market consolidation. In addition, the technology of protocols on which digital currencies applied to financial markets are based can cause numerous marginal innovations, enabling more secure, agile and customized products and gradually transforming the current structure.

3 BITCOIN

3.1 Characteristics

Developed by Satoshi Nakamoto (2008), possibly a pseudonym used by programmers, BTC was deployed on January 3, 2009 and is the most popular encrypted digital currency in the world. According to Nakamoto (2008), use of the P2P⁴ network enables operations without the strict need for third parties, transactions are recorded chronologically and information remains available.

The P2P architecture is the same as that used in file-exchange services such as the former *Kazaa* and current *Torrents*. P2P networks allow data transmission without the need for a central server and have the capacity to be self-organizing and fault-tolerant. BTCs are validated using algorithms. A *hash* function can map variable-length data and return fixed-length data. According to Paar and Pelzl (2009), *hash* is a widely used concept. It is a unique data identifier and must necessarily have a one-way function: because the *hash* is the data-entry product, it is impossible to obtain original data from it.

Each BTC user has two "keys": one public and one private. The public key identifies the user on

⁴ P2P-type networks are popularly known to facilitate the exchange of files using *torrents*. In this system, each machine works simultaneously as a server and user.

the network and uses a digital signature algorithm called ECDSA. Each user can have as many public keys as desired. For every transaction, the users involved must provide their respective public keys, which can be checked by any network user.

BTC issuance is limited to 21 million and the issue rate is decreasing over time. Taylor (2013) indicates that 58.8% of the total were issued by 2013 and by 2032, that percentage will be 99%. To date, as the chart below shows, approximately 14,357,900.00 BTCs have been mined. System transparency is a factor that makes BTC popular. At any time, it is possible to verify the amounts traded; however, it is not possible to access the direct identification of those involved in the transactions.

Another characteristic of BTC is *irreversibility*, i.e., the negotiator assumes all of the risk to a much greater degree than in a traditional monetary transaction. Records of transfers are permanently kept and according to Taylor (2013), these records could compromise anonymity, even for users who use multiple accounts, because a thorough analysis could reveal the same operator with a relative degree of certainty.

This possibility of checking transaction records made available by some traders to provide credibility to the market could also jeopardize the secrecy of BTC users. However, there are services that aim to make such evidence less clear. These use transmission instruments that are conditional upon future return, but involve a considerable number of accounts, thereby acting as a virtual scrambler.

BTC is widely used as a financial asset capable of generating speculative returns; however, they can also be useful as a tool to make foreign exchange transactions less costly (PAVEL, D'ARTIS and MIROSLAVA, 2014). Despite its popularity, several commercial networks are reluctant to use BTC because most governments prohibit the free movement of other currencies. Russia and China have already indicated that they intend to ban the use of digital currency. Singapore, Taiwan and the USA have expressed concern about illegal activities and defense of the national currency. In addition, the USA, Japan, Germany and Finland have created instruments to institute taxation.

May (1994) describes both governmental efforts to limit their populations' access to encryption tools and the adverse effects that these efforts have caused, for example, in France, the Philippines and the USA. The study's conclusions highlight the possibility of tax evasion because BTC can significantly compromise the ability of governments to collect taxes. According to Pavel, D'Artis and Miroslava (2014), BTC may be used as a safeguard in a turbulent market, as observed in the financial crisis of 2008, in which the adopted policy flooded the market with liquidity and exposed the adverse effects of an expansionist policy, among them, inflationary acceleration.

3.2 Electronic Currency characteristics

Electronic currency may come to represent a role similar to gold because its issue rate is not only known in advance but is also less than that of traditional currencies. In this case, the value of an electronic currency against other currencies, most likely even in non-turbulent times, depends on agents accepting that type of currency. There are three ways to acquire BTCs. At the beginning of its issue, the most common way to acquire was the *mining* process. At that time, there was not yet an established market. Nakamoto (2009) explains the term as analogous to gold mining activity. In the case of BTC, effort is represented by greater computational difficulty and electrical energy expenses.

BTC is mined in blocks. The first block mined was called the *Genesis block* and generated fifty BTCs. It should be noted that each hash consists of 64 characters and that each block mined informs the next block's hash. The success of mining generates a record in the blockchain, which is a file that records all transactions made in BTC. Each BTC found by a *miner* must be validated by other users in a systematic and algorithmic fashion.

The popularization of the currency, the increased number of miners and the increased computational power required have made mining less lucrative. In practice, users work in groups called *pools*, sharing the gains among group members based on their individual efforts. Indeed, explorers with little computational power are compelled to work in groups to obtain any return. The higher the processing power of a group of miners, the greater the likelihood of success. The degree of difficulty is

regulated independently so that it is possible to maintain a relatively constant decreasing issue rate.

The temporal dynamic creates a future with constantly decreasing issue rates, further complicating mining activity for new entrants. It is possible that by 2032 there may be fewer and more efficient miners, given the greater required computational power and higher hardware costs. In this sense, Taylor (2013) states that two factors are crucial: the first is access to energy and the second is having hardware that is paid off while the returns are high.

In the event of hoarding by miners, end users can develop a relative aversion to the currency. Even considering the extreme divisibility of BTC, users are subject to purely speculative currency attacks. However, there are two reasons that it is unlikely that such processes would be established: (1) it is impractical to coordinate many miners; and (2) miners do not have a matching entry point. Thus, the optimal price between miners differs considerably, even if they are a part of the same pool.

Depending on the entry point of the investment and operating costs, which are individual, if a large number of miners decide to take profits, the prices might never sink below than the cost of mining because at the moment that the market price is equal to the cost of mining, the miners who have not yet reached the point of exit would acquire BTC on the market and keep the machinery mining.

The mining process enables the miner both to be rewarded and to validate transactions. When a buyer of goods or services makes a payment order with BTC, the transaction is not carried out instantly. The buyer's digital wallet has a series of BTCs or fractions that were previously generated through computational effort. The goal of mining is to generate new blocks and these are generated after the "resolution" of a hash function.

Monetary transfers are made through a hash, which carries information from previous transactions and the public key of the recipient of the amount. When a miner deciphers the code and the code is validated by the network, the block is recorded in the blockchain and only after this process do the recipient, trader and miner receive the amount due. Because the block contains past information and the recipient's public key, all network nodes can be sure that the amount that the buyer wishes to pay still belongs to him. This eliminates the possibility of double spending.

The second way to obtain BTC involves free exchange between users. In general, this type of transaction incurs no monetary cost; as Taylor (2013) states, it is symbolic, ranging from 0.0005 BTC to 1 *Satoshi*, which is equivalent to 0.00000001 BTC. Each network node has a copy of the blockchain and because the public key used "assigns" a hash to each transaction, it is possible to compare the trader's public key, eliminating the possibility of double spending. In other words, if the trader no longer owns the traded BTCs, the stored history acknowledges this difference.

Verification of possession on the issuer's part and acceptance of the transfer by the recipient are stored in the blockchain and become available and updated information for each network node. It should be emphasized that the transaction will be stored in a new block, so mining activity is necessary because the transactions do not occur instantly. This modality involves BTC end users, traders and consumers of goods and finished products. For values to be transmitted, the user needs to use his private key.

Although this form of acquisition is the least expensive, it is also the riskiest. Those acting as buyers of a particular good or service have nobody to turn to if an agreement is breached. In reality, this is a determining factor for the currency's penetration capacity because for such a risk to be taken, the value of the good or service purchased would need to be relatively low. Otherwise, some form of guarantee is required.

This form of direct purchase of digital currencies makes it possible to render financial assistance to criminal causes and organizations. The possibility of those responsible being identified and punished are minimal if appropriate identification-protection strategies are used. Terrorist groups, paramilitaries, cartels interested in tax evasion, and even tacit agreements can be made explicit and undetected by traditional audit tools.

In contrast, humanitarian aid, donations to medical research and donations to those combating the excesses of governments that restrict individual freedoms can be considerably mitigated through the use of decentralized currencies. Nevertheless, it is important to note that when a big company decides to

adopt BTC, the currency begins to enjoy a credibility that is solidified over the time of its presence on the market, making direct negotiation considerably safer.

The third way to obtain BTC, and currently the most reported in the media, is the trading of intermediated currency. In this case, the user obtains the currency directly from various BTC "exchanges" or "banks." Meiklejohn et al. (2013) note that at the beginning of 2012, the percentage of transactions of less than the monetary unit abruptly doubled and those relating to less than one-tenth of the unit tripled. Not only has the trading of fractional amounts become more popular, but an increase in the velocity of the currency has also been observed⁵ and quantified; moreover, it is not an approximation. According to the above authors, only a portion estimated at four million of the available BTCs is traded frequently, but at a high exchange rate.

The spread of BTC is not uniform and many traders avoid adopting it because there are no clearly defined rules and there are issues relating to tax and the high volatility of the currency price. Keeping the currency over a long period can mean significant gains or losses to someone who transacts respecting the parity observed with his or her local currency at the time of purchase. In other words, the *carrying cost* increases, as mentioned previously. Despite users' desire for BTC and other digital currencies to be legalized and regulated, this initiative is at odds with the fundamental ideal of non-regulation and the assumption of all risks on the part of users.

Digital currencies provide growth opportunities for new niche markets known as *long-tail markets*. In these types of markets, the production of certain goods and services sometimes becomes uneconomical because of difficulties in reaching a geographically dispersed consumer market. In addition to other e-commerce tools, digital currencies may make such activities profitable by making exchange processes less bureaucratic and costly. This enables the trading of low-cost goods because there are practically no fees.

The World Economic Forum report (2015, p.102-108) presents three possible scenarios for the future role of financial institutions in response to continuous changes in the preferences of international financial system agents. In the first scenario, a market deconcentration movement is considered with the emergence of new agents and more flexible, intuitive and personalized financial products. In this scenario, current agents evolve and specialize, customizing their products to the greatest extent possible. Risks associated with this scenario include the possibility that the system will become hostage to technology developers, compromising the capacity for customer loyalty and creating uncertainty about regulatory aspects.

The second scenario presented depicts the possibility that the traditional banks shift focus to act as platforms for connecting niche products. In this case, there would also be a lower degree of loyalty and moreover, further specialization should occur, reducing the possibility of cross-subsidies. Critical points include the difficulty of selection because of greater atomicity, the difficulty of assigning responsibility for service failures or fraud, information-asymmetry problems and decreased customer loyalty.

Finally, in the third scenario, financial institutions would increase the points of contact with customers through technology, emulating the human relationship and aiming at retention and loyalty. The goal is to overcome needs with the provision of non-financial services without a significant increase in costs. It is thus possible that non-centralized and unregulated digital currencies will soon be regarded as a radical innovation, with disruptive innovation characteristics. In this case, incremental innovations may emerge, reshaping traditional forms of commerce, taxation and legislation.

All four types of innovation—commercial, organizational, technological and institutional—might occur to a greater or lesser degree depending on the degree of use of decentralized currencies. Institutional innovation is linked to the legally considered means of acting. In this sense, it is essential

⁵ Currency velocity is defined as the number of times that currency stock circulates between agents in a given time. The usual formula to estimate currency velocity considers production as a proxy for the number of transactions. However, the use of digital currencies, particularly BTC, means that the actual number of transactions is known, because each exchange is recorded.

that all aspects of a society's legal framework and regulatory institutions, including economic aspects, are indeed both effective and able to provide legal certainty to agents of various market types.

If BTC becomes an asset that is widely accepted as a currency, but the laws do not follow its evolution from the point of view of the legal framework applied to correct economic incentives, it will become largely worthless. Reference is made in this case to the entertainment industry's practices related to copyrights in some countries. Albuquerque and Callado (2015) recognize and provide examples showing that BTC regulation is weak, but despite regulatory issues, on May 19, 2015, the New York Stock Exchange (NYSE) created an official conversion rate between BTC and the US Dollar. This shows that the traditional financial market already accepts the digital currency market as a reality.

4 TWO-SIDED MARKET (2SM)

This section defines and characterizes the theory known as the Two-Sided Market (2SM). Despite the wide range of possible configurations, we will compare the structure of the electronic payment-card industry with the developing structure of the BTC market. Key similarities and differences that impede the BTC market's full conformity with the 2SM structure will be identified.

4.1 General Characteristics of the Two-Sided Market

Models that describe the 2SM have two characteristics: the existence of two distinct groups of participants and the presence of positive network externalities. This structure is based on a platform that allows contact and interaction between two types of user and that must be able to allow the groups to conduct the greatest possible number of transactions. The platform should be structured to promote the perception of the benefit of use on the part of end users.

Rochet and Tirole (2006) state that the 2SM is characterized not only by the existence of a platform that connects two end users but also by the fact that the total volume of transactions primarily depends on the price structure, i.e., how the price is divided between end users. According to Freitas (2007), in the 2SM, the pricing structure and definition of who will assume the majority of the platform costs must be "designed" in such a way that encourages participation by the two sides of the market. It follows that the price paid by a particular participant to use the platform is not necessarily related to the cost of entering the system.

Evans and Schmalensee (2005) cite some markets whose nature and characteristics result in their organization as 2SM, including dating agencies, night clubs, brokers, advertising agencies, newspapers and magazines, computer operating systems, video games, shopping centers and payment cards. According to the authors, this type of structure has three key features: differentiated prices, business design models and the existence of specific rules and regulations.

According to Evans (2003), final prices are charged to end users, consumers and traders to achieve demand balance. It is important both to maintain users and to attract new entrants from both sets of users to the network. The key element of this market, therefore, is balancing the demands of end users. The demand price elasticities for each group play a fundamental role in determining the respective prices charged to each of them. The price tends to be reduced on the more elastic side to facilitate users' entry into this side of the network. In general, the side with lower elasticity tends to bear the majority of the industry's costs.

Business design is crucial in the analysis of the 2SM not only because of the presence of network externalities and demand price elasticities but also because of the need to integrate the two market sides. The existence of specific rules and regulations imposed by the platform is necessary because of the need to construct parameters that promote positive externalities and limit negative ones to produce benefits for end users (BANCO CENTRAL DO BRASIL; SECRETARIA DE ACOMPANHAMENTO ECONÔMICO; SECRETARIA DE DIREITO ECONÔMICO, 2010).

The pricing structure should promote incentives for at least one side without resulting in a very severe punishment to the other, thereby discouraging its participation. This price-fixing process in two-sided markets reflects not only the industry participants' cost structure but also the surplus generated for

one of the groups when the market incorporates one user more than another (ROCHET and TIROLE, 2002; 2003). This means that in the pricing structure of a 2SM, if all other conditions remain constant the group that obtained the largest surplus *subsidizes* the other group. The side that bears the costs is called the profit center and the side that benefits is called the loss center.

4.2 Characteristics of the Card Market

In the analysis of the 2SM pricing structure, two characteristics should be observed. The first is the sum of prices coming from the two-sided structure and the second is the distribution of these prices among end users on either side. The prices on each side of the market depend not only on their marginal cost but also on the demand price elasticity of each side of the market, on the value that the entry of an agent on one side generates for participants on the other side (network externality), and on the degree of inter-platform and intra-platform competition (BANCO CENTRAL DO BRASIL; SECRETARIA DE ACOMPANHAMENTO ECONÔMICO; SECRETARIA DE DIREITO ECONÔMICO, 2010).

Fagundes, Ferrés and Saito (2009)⁶ explain that externalities of use arise with increases in revenues because of diffusion of the tool. The internalization of positive externalities for trade and consumers can increase the volume of transactions and consequently increase revenues if the burden of use is primarily borne by the agent with less elasticity.

In the traditional case of this type of market structure, the attempt to make the market more competitive based on regulations supported by technical studies that are strongly based on microeconomic theory and that consider price convergence at a level close to the respective marginal costs could eliminate the characteristic of network externality and produce effects of greater magnitude than the surplus generated by the regulatory body. For example, a higher level of competition can only result in lower profitability for the owner. However, the literature⁷ suggests that stimuli for competition can promote benefits when a single user has access to multiple cards indiscriminately accepted by merchants.

It could be argued that regardless of the different elasticities among end users, the cost could be passed on (either in whole or in part) from traders to consumers. In other words, by presenting different costs for different types of payments, traders can differentiate and pass on those costs to customers; however, preclusion clauses are imposed on trade groups that would make it difficult to differentiate prices. However, the inability to differentiate prices could cause cross-subsidy, punishing users who choose the less expensive form of payment.

The card market is frequently structured in the form of a platform containing three or four parts. In the three-part platform, the accreditor is also the card issuer. In the four-part platform, these agents are not the same.

Briefly, the market works as follows: the scheme owner (card brand) defines the business rules; the issuers are usually banks that maintain direct contact with cardholders and define credit limits, fees and charges, invoice due dates, benefits programs, etc.; the accreditor then provides a bridge between commercial establishments and the scheme owner by contract; the consumer is the cardholder; the seller is the agent who accepts the card as a means of payment; and the final prices are shared among end users, consumers and sellers to balance the demands of both sides.

5 METHODOLOGICAL PROCEDURES

This section presents the technical procedures involved in this study's empirical strategy. The GARCH (generalized autoregressive conditional heteroskedasticity) family models will be described briefly to help understand BTC's volatility. As mentioned above, this study was based on the contributions of Pavel, D'Artis and Miroslava (2014) and Stråle and Tjernström (2014). In the former case, vector auto-regressive (VAR) models were estimated with variables that are also used in this study. In the latter study, the authors estimated a GARCH model to show that demand, volume traded, the Mt

⁶ Although the authors' focus was on the credit card market, similar reasoning applies to the object of this study, BTCs.

⁷ See Evans (2002), Frascaroli (2010) and Rochet and Tirole (2003) for more on this issue.

Gox bankruptcy and the Cyprus crisis, which resulted in capital controls in mid-March 2013, were determinants of BTC price volatility.

Here, estimations of the DCC MGARCH family models aimed to estimate the parameters of changes in volatilities and their quasi-correlations between variables. The SPX-500, the price of gold and China's main stock index (SSEC) were considered for this purpose. Volatility is a characteristic often present in financial asset series quantified by the standard deviation (or variance) of results. Used as an asset risk measure, volatility can be measured by the intensity of variations in the historical process analyzed, which can behave unpredictably based on many factors, including institutional factors, legal factors, macroeconomic factors, or aspects more closely linked to the formation of the market in which the asset in question, in this case, BTC, is negotiated.

To capture the presence of volatility in financial series, it is necessary to use models whose estimators are adapted for historical processes of a heteroskedastic nature. These models consider the variance (volatility) of a return at a given point in time as dependent on past information and therefore consider it as the historical process of a random variable that follows a stochastic process. Engle's seminal work (1982) introduced the possibility of estimating autoregressive conditional heteroskedastic models (ARCH).

5.1 GARCH Multivariate Model (MGARCH)

Engle (1982), to estimate the variance of US inflation, shows that linear models are limited and unable to explain some features. He therefore shows that it is possible to capture the volatility of autocorrelations. The generalized autoregressive conditional heteroskedasticity (GARCH) model, formulated by Bollerslev (1986), is a generalization of the ARCH model. It allows the use of only two parameters to simultaneously capture the mean and variance of the time series of an ARMA process. A GARCH (r, s) model, in which parameter r represents the order of the ARCH component and s the order of the GARCH component. Following the seminal work of Engle (1982), some extensions of the ARCH model were developed, including GARCH, EGARCH and TARCH.

The basic idea stems from the understanding of the existence of heteroskedasticity conditional on the immediate past observation, in other words, that the present return depends on the past return. MGARCH allows the estimation of generalized autoregressive parameters considering conditional heteroskedasticity. Because they are dynamic, MGARCH models allow the mean and conditional covariance to be obtained. This model would be too flexible and it would not be feasible to estimate all parameters. For this reason, a wide variety of GARCH multivariate models enable parameters to be obtained in a more parsimonious manner.

In general, four parameterization methods are used alternatively: the Diagonal VEC model (DVECH), the constant conditional correlation model (CCC), the dynamic conditional correlation model (DCC) and the time-varying conditional correlation model (VCC). Bollerslev, Engle and Woodridge (1988), Bollerslev, Engle and Nelson (1994) and Bauwens, Laurent, and Rombouts (2006) describe general aspects of MGARCH modeling. A definition that allows comparison between alternative models is given by the following equation:

$$y_t = Cx_t + \epsilon_t \quad (1)$$

$$\epsilon_t = H_t^{1/2} v_t \quad (2)$$

where y_t is a dependent variable vector of the order $m \times 1$; C is a parameter matrix of order $m \times k$; x_t is a vector of independent variables that may contain lags of y_t ; $H_t^{1/2}$ is the Cholesky factor of the conditional time-varying covariance matrix H_t ; and v_t , which is a vector of order $m \times 1$ is with zero mean and i.i.d., with variance equals to unit.

In the general model, H_t is a matrix of univariate GARCHs. For example, a general MGARCH with an autoregressive term of conditional heteroskedasticity (ARCH) and a GARCH term is given by the following equation:

$$vech(H_t) = s + Avech(\epsilon_{t-1}\epsilon'_{t-1} + Bvech(H_{t-1})) \quad (3)$$

where the $\text{vech}(\cdot)$ function captures elements above or below the main diagonal; s is a parameter vector and A and B are parameter matrices. Because this model uses the vech function to extract and model unique elements of H_t , it became known as the VECHE model. H_t must therefore be positively defined. Equation (3) can be used to show that the parameters in s , A and B are not unique; thus, for H_t to be positively defined other restrictions should be incorporated into s , A and B . The models that the literature proposes vary with respect to the degree of flexibility and parsimony of the specifications imposed on H_t .

5.1.1 MGARCH Diagonal vech (DVECH)

Bollerslev, Engle, and Wooldridge (1988) require A and B to be diagonals. Although more parsimonious than the general case, the model is only advantageous when estimating a small number of temporal processes because the number of parameters increases quadratically. Despite the large number of parameters, the diagonal structure implies that each conditional variance and covariance depends on their own past values, not on other variances and covariances. Formally, in DVECH (1,1) each element of H_t is given as follows:

$$h_{ij,t} = s_{ij} + a_{ij}\epsilon_{i,(t-1)}\epsilon_{j,(t-1)} + b_{ij}h_{ij,(t-1)} \quad (4)$$

The process requires H_t to be defined as positive in each t , which imposes severe restrictions.

5.1.2 MGARCH Conditional Correlation Models (CC)

In these models, a non-linear combination of GARCH univariate models is used to represent conditional variances. The conditional correlation matrix is defined positively, motivated by the building structure that facilitates the estimation of parameters. In MGARCH conditional correlation models (CC), H_t is decomposed in a conditional correlation matrix R_t and a diagonal matrix of conditional variances D_t :

$$H_t = D_t^{1/2} R_t D_t^{1/2} \quad (5)$$

where each conditional variance follows a univariate GARCH process and the parameterizations of R_t vary according to the specific model. According to (5) we have

$$h_{ij,t} = \rho_{ij,t} + \sigma_{i,t}\sigma_{j,t} \quad (6)$$

where $\sigma_{i,t}$ is modeled by a univariate GARCH process. The equation highlights that conditional correlation models incorporate non-linear combinations of GARCH univariate models to represent conditional variances and that the parameters in the model $\rho_{ij,t}$ describe the joint dynamics of errors observed in equations i and j . Comparing (4) and (5), it is clear that in conditional correlation models, the number of parameters increases more slowly.

Bollerslev (1990) proposes a conditional correlation model with a time-invariant correlation matrix. For this reason, the model is known as the MGARCH constant conditional correlation model (CCC). By restricting R_t to be constant, the number of parameters reduces and simplifies the estimation; however, this fact can significantly restrict the accuracy of estimation and compromise the analysis.

5.1.3 Dynamic Conditional Correlation Models (DCC)

Engle (2002) introduced a dynamic conditional correlation model (DCC) wherein R_t follows a GARCH (1,1) process. As described by Engle (2009), the parameters of R_t are not standardized to be correlations and are called quasi-correlations. To maintain parsimony, all conditional quasi-correlations are constrained to follow the same dynamics. This model is significantly more flexible than the CCC and does not introduce a much larger number of parameters for the number of series. Because of this advantage, DCC MGARCH modeling was used for this study.

This model uses weighted non-linear combinations of time-variant univariate GARCHs. The diagonal elements of H_t are modeled using univariate estimates (GARCH) and the remaining elements are modeled by nonlinear functions based on the terms resulting from the diagonal. In a DCC MGARCH we have

$$(12)$$

where the terms $h_{ii,t}$ and $h_{jj,t}$ follow the GARCH process and $\rho_{ij,t}$ follows the dynamic process as outlined in Engle (2002). The DCC MGARCH can be written as follows:

$$y_t = Cx_t + \epsilon_t \quad (7)$$

$$\epsilon_t = H_t^{1/2} v_t \quad (8)$$

$$H_t = D_t^{1/2} R_t D_t^{1/2} \quad (9)$$

$$R_t = \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2} \quad (10)$$

$$Q_t = (1 - \lambda_1 - \lambda_2)R + \lambda_1 \tilde{\epsilon}_{t-1} \tilde{\epsilon}_{t-1}' + \lambda_2 Q_{t-1} \quad (11)$$

where the only term still unexpressed is D_t , then we have that y_t is a dependent variable vector of the order $m \times 1$; C is a parameter matrix of the order $m \times k$; x_t is an independent variable vector that may contain lags of y_t ; $H_t^{1/2}$ is the Cholesky factor of the time-variant conditional covariance matrix H_t ; v_t is a vector of the order $m \times 1$, which is a vector of zero mean and i.i.d., with variance equals to unit; and D_t is a diagonal conditional variance matrix.

$$D_t = \begin{bmatrix} \sigma_{1,t}^2 & 0 & \dots & 0 \\ 0 & \sigma_{2,t}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{m,t}^2 \end{bmatrix}$$

where each $\sigma_{i,t}^2$ evolves according to a univariate GARCH given as follows:

$$\sigma_{i,t}^2 = s + \sum_{j=1}^{p_i} \alpha_j \epsilon_{i,t-j}^2 + \sum_{j=1}^{q_i} \beta_j \sigma_{i,t-j}^2 \quad (12)$$

or when the constant term is included:

$$\sigma_{i,t}^2 = \exp(\gamma_i z_{i,t}) + \sum_{j=1}^{p_i} \alpha_j \epsilon_{i,t-j}^2 + \sum_{j=1}^{q_i} \beta_j \sigma_{i,t-j}^2 \quad (13)$$

where γ_i is a $1 \times p$ parameter vector, z_i is a $p \times 1$ vector of independent variables including a constant term, α_j are ARCH parameters and β_j are GARCH parameters. The conditional quasi-correlation matrix will be:

$$R_t = \begin{bmatrix} 1 & \rho_{12,t} & \dots & \rho_{1m,t} \\ \rho_{12,t} & 1 & \dots & \rho_{2m,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1m,t} & \rho_{2m,t} & \dots & 1 \end{bmatrix}$$

where $\tilde{\epsilon}_{t-1}$ is a $m \times 1$ vector of standardized residuals, $D_t^{-1/2} \tilde{\epsilon}_t$; and λ_1 and λ_2 are parameters that govern the quasi-correlation dynamics. In addition, λ_1 and λ_2 are non-negative and meet condition $0 \leq \lambda_1 + \lambda_2 < 1$. When Q_t is stationary, matrix R is the weighted average of the covariance matrix of the standardized residuals $\tilde{\epsilon}_t$, denoted by $\bar{R}$ and the unconditional average of Q_t and denoted by $\bar{Q}_t$. Because $Q_t \neq R_t$, as demonstrated by Engle (2009), R does not represent either the correlation matrix or the unconditional average of Q_t ; for this reason, the parameters in R are called quasi-correlations.

6 SAMPLE PLANNING AND TREATMENT

All of the data used in the study were obtained from the site *br.investing.com*. A sample of 1,362 observations was used and the period from September 13, 2011, to June 23, 2015, was considered. In addition, to test the robustness of the estimates, a subsample was built, which considered only the data from March 10, 2013, and on, with 1,192 observations. The latter is explained by events and trends that impacted the historical series of BTC prices, as discussed in the text.

One of the technical instruments justifying the need to analyze the samples separately was MS-VAR modeling, which indicated the convergence of the historical series of variables considering two regimes, with regime 1 showing an expected duration of 75% of the time period and regime 2 having an expected duration of 25%. This model was used to define the time window for determining the subsample, i.e., in addition to visual and historical evidence, the MS-VAR of the two regimes was estimated. Furthermore, the ARCH and GARCH effects, as demonstrated by Stråle and Tjernström (2014), were evident.

For analytical purposes at the microeconomic level, the 2SM theory was used to explain one of the possible phenomena associated with this phase, and a second phase of “convergence” to the fundamentals that determine the prices of the currency. The last one persists until the end of the series, where peaks of the same magnitude are no longer observed.

At level, the series presents undesirable characteristics; an abyssal range is observed between the maximum value of US \$1,132.01 and the minimum value of \$ 2.24. By visually comparing the sample returns to the subsample returns, the differences in magnitude are clear. Series denoted in the form of returns enable the mitigation of persistent trend effects; however, visual analysis suggests periods of high volatility that possibly affect variance over time. According to this analysis, GARCH family models are suitable to attempt to quantify conditional variance effects.

In this study, the hypothesis of normality was rejected for all variables in all samples. In addition, excess kurtosis indicates leptokurtic series that are common in financial time series, thus strengthening the argument for the use of models capable of measuring volatilities and studying how they are correlated. In addition to the BTC price, the SPX-500, the price of gold and China's main stock index (SSEC) were used. In all cases, the series showed both skewness and kurtosis. Descriptive statistics of the sample and subsample of variables used are shown in Table 1:

Table 1 - Descriptive statistics of the sample* and subsample**

Variables	Mean	Median	Maximum	Minimum	Standard Deviation	Skewness	Kurtosis
<i>BTC*</i>	0.001256	0.000389	0.450252	-0.409647	0.0293	1.245677	89.84125
<i>BTC**</i>	0.000483	0.000211	0.066043	-0.130503	0.011374	-1.639805	30.0825
<i>XAU*</i>	-0.0000439	-0.00000638	0.006592	-0.013292	0.001278	-1.129933	15.68017
<i>XAU**</i>	-0.0000488	-0.0000198	0.006592	-0.013292	0.001313	-1.356306	19.10513
<i>SSEC*</i>	0.0000565	0.000021	0.00575	-0.007908	0.001198	-0.348153	9.388075
<i>SSEC**</i>	0.000106	0.000103	0.00575	-0.007908	0.001223	-0.762647	10.46889
<i>SPX-500*</i>	0.0000597	0.0000322	0.005984	-0.007674	0.000924	-0.43433	11.42462
<i>SPX-500**</i>	0.0000503	0.0000413	0.00312	-0.003425	0.000747	-0.186709	5.734772

Source: The authors' work based on selected data.

7 RESULTS

This section presents all of the statistical procedures for the estimates of volatility measures according to the models presented in section 5. The series is sensitive to extreme values, but these observations should not be suppressed because the size of the oscillation and the existence of wide gaps characterize the BTC returns series. As described above, a visual analysis of the BTC series revealed that it showed a tendency to appreciate in times of economic and political crisis.

Moreover, attacks on financial institutions result in the opposite tendency. Despite the impossibility of predicting economic cycles deterministically, an analysis of macroeconomic fundamentals can provide BTC speculators with information that they can use to change strategy. Crises seem to act as an expansion factor of both the BTC market and digital currencies in general. In terms of technical analysis of the data, it is reported that the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests indicated stationarity of the returns series, as shown in Table 2. It therefore follows that like the other variables, BTC returns are stationary and therefore do not depend on time.

Table 2 - Unit Root Tests

Unit Root Tests (Sample)			Unit Root Tests (Subsample)		
Variables	ADF	PP	Variables	ADF	PP
<i>BTC</i>	-8.301624*	-41.44987*	<i>BTC</i>	-28.14421*	-28.16805*
<i>XAU</i>	-32.75358*	-32.82711*	<i>XAU</i>	-29.23710*	-29.24580*
<i>SSEC</i>	-36.55772*	-37.03559*	<i>SSEC</i>	-23.68373*	-23.76906*
<i>SPX-500</i>	-37.20030*	-37.20349*	<i>SPX-500</i>	-28.87981*	-28.87969*

Source: The authors' work based on selected data. Note: *The null hypothesis is rejected at 1%.

The choice of best model by degree of fit was based on the Akaike Information Criterion (AIC) and the Schwartz Bayesian Criterion (SBC). According to Tsay (2005), the smaller the AIC and SBC, the better the model fit to the data series. Using the parsimony criterion, the ARMA model (0,2) offered the

best fit and had the expected characteristics of non-correlated returns and severe correlation of the squares of returns, which justifies estimation by models capable of capturing conditional volatility effects.

Table 3 - Best fit ARMA model

Model	Schwarz criterion	Akaike Criterion
<i>BTC AR(0) MA(2) Sample</i>	-4.31091	-4.318593
<i>BTC AR(0) MA(2) Subsample</i>	-6.108152	-6.101125

Source: The authors' work based on selected data.

The third step in building the ARCH model was to attempt to fit the ARMA model using the *sample* to verify the serial correlation of the series. It should be noted that the data have a long memory; thus, the ARCH test only captured volatility from the second residual lag onwards. Table 4 shows the test results:

Table 4 - F Statistic - ARMA (0,2)

Test: ARCH - ARMA Model (0,2)			
F Statistic - Sample	8.098550	Prob. F (4.1348)	0
Obs* R ²	31.75133	Prob. Chi-square (4)	0
F Statistic - Subsample	52.71409	Prob. F (4.819)	0
Obs* R ²	168.7086	Prob. Chi-square (4)	0

Source: The authors' work based on selected data. *The series shows volatility

Among the estimated models, the best fit model for the *sample*, with 1357 observations, was the GARCH (1,1) according to the Schwartz criterion. The model completely corrected residual autocorrelation and when the ARCH test was repeated, a lack of conditional volatility was found. The resulting estimated model is given by

$$h_t = (3,46)10^{-6} + (0,220539)X_{t-1}^2 + (0,800577)h_{t-1} \quad (8)$$

Tables 5 and 6 summarize the results of the model and its F statistic, respectively:

Table 5 - Results of GARCH Model (1,1)

Variable	Coefficients	Standard Error	Z statistic	P Value
<i>Intercept</i>	0.000114636	0.00029024	0.39497012	0.69286494
<i>MA (2)</i>	0.055499339	0.027560405	2.013734487	0.044037417
Variance equation				
<i>Intercept</i>	3.46E-06	2.17e-07	15.93489162	3.63E-57
<i>Residuals (-1)²</i>	0.220539029	0.012605408	17.49558823	1.55E-68
<i>GARCH (-1)</i>	0.800576674	0.008196231	97.67619844	0
<i>R²</i>	-0.037313333			
<i>R² adjusted</i>	-0.038078878			
<i>SQR</i>	1.207575369		<i>Akaike criterion</i>	-6.039195489
<i>DW statistic</i>	2.103768306		<i>Schwarz criterion</i>	-6.019987561

Source: The authors' work based on selected data.

Table 6 - F statistic – GARCH (1,1)

Test: ARCH GARCH Model (1,1)			
F statistic	0.004481	Prob, F(1.1354)	0.9466
Obs* R ²	0.004488	Prob, Chi-square (4)	0.9466

Source: The authors' work based on selected data.

Thereafter, autocorrelation tests of the returns residuals and squared returns residuals—i.e., the *Q* statistic—were performed to dispel doubts about the robustness of the results. Table 7 shows the test results:

Table 7 - Autocorrelation test for residuals and the square of residuals— Q Statistic

Autocorrelation	Partial autocorrelation	t	AC	PAC	Q-Stat	Prob*
		1	0.047759972	0.047759972	3.102185399	
		3	0.032210437	0.034184226	4.973127558	0.083195354
...	...	...	...	...	...	...
* ¹	*	7	0.097656629	0.093272095	22.84401255	0.000850442
		1	0.001819164	0.001819164	0.004500734	0.946512036
		2	0.006343791	0.006340502	0.059272623	0.970798538
* ²	*	3	-0.023150262	-0.023174272	0.789221532	0.852043326

Source: The authors' work based on selected data. ¹Residuals. ² Square of residuals.

The relevant analysis rests on the variance equation coefficients, which were all significant at the 1% level. With $\alpha_1 = 0.220539$ and $\beta_1 = 0.800577$ it is clear that both immediate shock and persistence are quite representative, but there is a preponderance of persistence.

Estimations similar to the previous ones were made for the 1,192 *subsample* observations and as expected, the model with best fit for volatility diverged. The ARMA model (0,2) also resulted in both a better fit according to the Schwarz criterion and the best fit for volatility for a GARCH model (1,1), i.e., an identical structure to that observed previously. However, the resulting coefficients are essentially different. The adjusted model is given as follows:

$$h_t = (6,74)10^{-6} + (0.240146)X_{t-1}^2 + (0.698707)h_{t-1} \quad (9)$$

Tables 8 and 9 summarize the results of the model and its F statistic, respectively:

Table 8 - Results of GARCH Model (1,1)

Variable	Coefficients	Standard error	Z statistic	P Value
Intercept	-0.000348	0.000286	-1.214364	0.2246
MA (2)	-0.010033	0.041015	-0.244615	0.8068
Variance equation				
Intercept	6.74E-06	6.30E-07	10.71209	0.0000
Residuals (-1) ²	0.240146	0.024194	9.926044	0.0000
GARCH (-1)	0.698707	0.021892	31.91609	0.0000
R ²	-0.04835			
R ² adjusted	-0.006051			
SQR	0.107511		Akaike criterion	-6.805601
DW statistic	1.948503		Schwarz criterion	-6.794671

Source: The authors' work based on selected data.

Table 9 - F statistic – GARCH (1,1)

Test: ARCH GARCH Model (1,1)			
F statistic	0.333927	Prob. F(1.825)	0.5635
Obs* R ²	0.334602	Prob. Chi-square (4)	0.5630

Source: The authors' work based on selected data.

As in the case of estimations for the *sample*, autocorrelation tests of returns residuals and squares of returns residuals—i.e., the Q statistic—were performed with data from the *subsample* to dispel doubts about the robustness of the results. Table 10 shows the results of this test:

Table 10 - Autocorrelation tests of residuals* and residuals²** – Q statistic

Autocorrelation	Partial autocorrelation	t	AC	PAC	Q-Stat	Prob*
*		1	0.063624561	0.063624561	3.363973072	0.024955675
		2	0.037853145	0.033942462	4.556126953	0.032801171
		3	0.027915075	0.023551218	5.205257526	0.074078587
...	...	...	...	...	...	...
*	*	6	0.11480386	0.108584665	18.26474391	0.002632269
* **	*	1	0.02011261	0.02011261	0.336155164	0.562056987

Source: The authors' work based on selected data.

In addition, the conditional variance model GARCH (1,1) eliminated the serial correlation observed in the ARMA modeling (0,2) and the ARCH test indicates the absence of the effect after proper estimation. When comparing this result with the previous estimation, it can be seen that the constant term has almost doubled, which may indicate the perception of a market that is not incipient. The parameter β_1 , however, lost relative importance, indicating that volatility persistence had declined. The term α_1 essentially did not change.

As described among the objectives of this study, the DCC MGARCH model (1,1) was estimated. Estimates using the *sample* indicated a series of volatile behaviors with an almost perfect degree of volatility persistence. Table 11 shows the results of the model best fit for the *sample*, and the model tests are shown in Tables 12 and 13:

Table 11 - MGARCH (1,1) model results

Parameter	Value	Standard error	t	t-prob.
ρ_{21}	0.074760	0.075238	0.9936	0.3206
ρ_{31}	0.002272	0.047111	0.04823	0.9615
ρ_{41}	0.016071	0.056024	0.2869	0.7743
ρ_{32}	-0.001535	0.070378	-0.02181	0.9826
ρ_{42}	0.068042	0.059743	1.139	0.2550
ρ_{43}	0.080281	0.065381	1.228	0.2197
$\square$	0.004629	0.0024437	1.894	0.0584
$\square$	0.989691	0.0083055	119.2	0.0000
Akaike	-38.659429		Shibata	-38.660821
Schwarz	-38.522963		Hannan-Quinn	-38.608006

Source: The authors' work based on selected data.

Table 12 - Normality tests for the MGARCH model (1,1)

	Statistic	Value	t-Test	p-Value
BTC	Skewness	-1.2121	17.105	1.3533e-065
	Excess kurtosis	11.351	80.163	0.00000
	Jarque-Bera	6691.2	-	0.00000
XAU	Skewness	0.20592	2.9061	0.0036594
	Excess kurtosis	4.0215	28.401	1.9845e-177
	Jarque-Bera	811.66	-	5.6319e-177
SSEC	Skewness	-0.38843	5.4818	4.2112e-008
	Excess kurtosis	3.6039	25.451	6.8162e-143
	Jarque-Bera	675.05	-	2.6045e-147
SPX500	Skewness	-0.57573	8.1251	4.4700e-016
	Excess kurtosis	5.6941	40.213	0.00000
	Jarque-Bera	1676.2	-	0.00000

Source: The authors' work based on selected data.

Table 13 - Autocorrelation tests of residuals* and residuals^{2**} – Q statistic

TC	(Q5) = 0.952249	[0.9663245]	SSEC	(Q5) = 2.31352	[0.8042793]
	(Q10) = 8.09680	[0.6193817]		(Q10) = 6.61360	[0.7613499]
	(Q20) = 13.3655	[0.8611587]		(Q20) = 13.3655	[0.8611587]
	(Q50) = 58.2710	[0.1973003]		SSEC ² (Q50) = 39.4433	[0.8583046]
BTC ²	(Q5) = 1.34048	[0.9307118]		(Q5) = 17.1297	[0.0042603]
	(Q10) = 4.41971	[0.9264334]		(Q10) = 45.6557	[0.0000017]
	(Q50) = 30.0825	[0.9884888]			
XAU	(Q 5) = 10.1627	[0.0707548]	SPX500	(Q 5) = 3.51488	[0.6211372]
	(Q 10) = 19.6281	[0.0329728]		(Q 10) = 9.58548	[0.4775813]
	(Q 50) = 48.5299	[0.5325121]		(Q 50) = 37.5876	[0.9021283]
XAU ²	(Q 5) = 10.3517	[0.0658617]	SPX500 ²	(Q 5) = 14.2121	[0.0143168]
	(Q 10) = 25.1507	[0.0050671]		(Q 10) = 17.8319	[0.0578655]
	(Q 50) = 102.662	[0.0000168]		(Q 50) = 72.4295	[0.0207066]

Source: The authors' work based on selected data.

According to analysis of the table 11 parameters using data from the *sample*, the degree of interdependence between the volatilities of the BTC returns series and the other variables was low. These results are in line with previous empirical studies such as Pavel, D'Artis and Miroslava (2014) and Stråle and Tjernström (2014). Their results did not indicate significant conditional variance coefficients for any of the variables considered.

The characteristics that hinder good model fit are revealed in the normality and serial autocorrelation tests shown in tables 12 and 13. For the latter, the null hypothesis that there is no serial correlation cannot be rejected because the probabilities $Q < \text{Chi squared}$ for all lags are high. Using the same specifications as before, a model was estimated for the *subsample* with 663 observations. Table 14 presents the best fit result:

Table 14 – Results of MGARCH (1,1) model

Parameter	Value	Standard error	<i>t</i>	<i>t-prob.</i>
ρ_{21}	-0.021574	0.045540	-0.4737	0.6358
ρ_{31}	0.062535	0.039347	1.589	0.1125
ρ_{41}	-0.074188	0.047476	-1.563	0.1186
ρ_{32}	-0.005211	0.052837	-0.09863	0.9215
ρ_{42}	-0.064615	0.054152	-1.193	0.2332
ρ_{43}	-0.008687	0.051871	-0.1675	0.8671
α	0.020865	0.024322	0.8579	0.3913
β	0.664926	0.27447	2.423	0.0157

Source: The authors' work based on selected data.

Observations: 663 Variables: 1-BTC; 2-XAU; 3-SSEC; 4-SPX500.

In this study, it was decided not to include binary variables in the model estimations as because would be impossible to keep them when estimating the multivariate models. To obtain a date limit for determining the subsample beyond the visual and historical evidence, the MS-VAR model of the two regimes was estimated. Furthermore, the ARCH and GARCH effects, as demonstrated by Stråle and Tjernström (2014), were obvious.

Some other changes in relation to the previous results should be highlighted. The variable α , which indicates the effect of the impact of volatility at the time the shock occurs, was no longer significant in the multivariate results. However, the SSEC variable became significant at 11.25% and the same occurred with the SPX-500 variable at 11.86%. Although above the 10% threshold, this convergence seems to indicate that the BTC markets are transforming and gradually becoming more interdependent by observing the estimated parameters. Tables 15 and 16 show the normality and autocorrelation tests, respectively:

Table 15 - Normality tests MGARCH model (1,1)

	Statistic	Value	t-Test	p-Value
BTC	Skewness	-1.1073	11.666	1.8963e-031
	Excess kurtosis	8.6850	45.819	0.00000
	Jarque-Bera	2219.2	-	0.00000
XAU	Skewness	-0.49425	5.2073	1.9163e-007
	Excess kurtosis	5.6201	29.650	3.3973e-193
	Jarque-Bera	899.55	-	4.6204e-196
SSEC	Skewness	-0.086449	0.91080	0.36240
	Excess kurtosis	3.1366	16.548	1.6633e-061
	Jarque-Bera	272.61	-	6.3640e-060
SPX500	Skewness	-0.25520	2.6887	0.0071723
	Excess kurtosis	2.7393	14.451	2.4555e-047
	Jarque-Bera	214.48	-	2.6658e-047

Source: The authors' work based on selected data.

Table 16 – Autocorrelation tests of residuals* and residuals^{2**} – *Q* statistic

BTC	(Q5) = 4.18133 [0.5236152]	SSEC	(Q 5) = 6.80759 [0.2353482]
	(Q10) = 12.7910 [0.2355925]		(Q10) = 18.1040 [0.0532340]
	(Q50) = 68.3933 [0.0429550]		(Q50) = 52.0189 [0.3952199]
BTC ²	(Q5) = 0.616186 [0.9872502]	SSEC ²	(Q5) = 9.04407 [0.1073194]
	(Q10) = 1.95980 [0.9966391]		(Q10) = 27.9164 [0.0018617]
	(Q50) = 31.8304 [0.9788777]		(Q50) = 111.270 [0.0000015]
XAU	(Q5) = 1.45537 [0.9181620]	SPX500	(Q5) = 2.16674 [0.8256249]
	(Q10) = 3.47511 [0.9679367]		(Q10) = 3.75388 [0.9577674]
	(Q50) = 40.8745 [0.8178095]		(Q50) = 60.1708 [0.1536360]
XAU ²	(Q5) = 12.4954 [0.0285950]	SPX500 ²	(Q5) = 11.0774 [0.0498660]
	(Q10) = 14.3132 [0.1591778]		(Q10) = 33.9054 [0.0001916]
	(Q50) = 47.4270 [0.5772369]		(Q50) = 168.501 [0.0000000]

Source: The authors' work based on selected data.

In summary, gold does not show statistical significance in any of the multivariate models estimated. Thus, these results possibly capture the inertial aspect of the "euphoria" phase. It should be recalled that at the beginning of the sample period, the currency was not used mainly as a medium of exchange because of the factors presented in relation to adaptation of BTC in relation to 2SM. It is therefore possible that most agents demanded BTC for the purposes of speculation, and perceived all negative returns as a buying opportunity, not as a reversal of trend in prices.

In addition, note that because of the many robustness exercises needed to achieve the models presented here, the profile of the BTC returns series renders hedging strategies virtually impossible for those who use it to invest, which increases the carrying cost for those who transact with the currency. In general, as discussed, it is possible that the series is dominated by shocks arising from expectations, whether because of rumors about prohibition or regulation or hacker attacks on active financial institutions.

The results obtained in the estimation of the DCC MGARCH model, although not significant, represent some progress because of the methodology used. It should be noted that Pavel, D'artis and Miroslava (2014) attempt to identify the BTC price formation process by VAR modeling and similarly, the parameters obtained by those authors are not significant.

This multivariate analysis could indicate volatility as a BTC price formation component. Although the persistence of autocorrelation might have affected the estimated parameters of the multivariate model for the *sample* and *subsample*, one can see that the pressure of a sudden increase in demand caused by political and economic crises and the observed effects of hacker attacks on the BTC exchanges seem to affect the estimated volatilities. Nonetheless, the results of the multivariate analysis for the most recent observations indicate convergence to significant parameters that seem to reflect the financial system's adherence to BTC. The perfect adaptation to the 2SM structure of debit cards may indeed be only one of many possible effects of this convergence.

8 CONCLUSIONS

To achieve the overall objective, the institutional and operational aspects of BTC were analyzed. The study revealed that the emergence of BTC coincided with a period of intense effects of the global crisis originating in the US. This historical and social context might have been decisive in BTC achieving popularity. In other words, it can be said that crises seem to act as an expansion factor of the BTC market and digital currencies in general.

The study revealed the existence of antagonisms limiting BTC's operation. It was found that the most restrictive laws are found in places in which the national currency has weak desirability and BTC is thus characterized as radical innovation because it does not conform to the structures observed before its emergence. When considering the conceptual typifications of disruptive innovation, BTC has potential characteristics of the *low disruption* and *new market* types, but its impact on the economic environment is fundamentally dependent on it being perceived as a currency. BTC and other non-centralized issue

currencies are apparently being absorbed by the financial system more intensely not as a currency, but as a financial asset.

The wide range of digital currencies available promotes competition, and the population might or might not choose a dominant type. The adoption of BTC by financial institutions, despite its potential for creative destruction, contrasts with the low number of traders willing to accept it. Perhaps financial institutions are not overly rigid in adapting their services. Digital currencies could represent opportunities for net gains when relativized with possible losses by enabling the supply of new products and the acquisition of new customers who did not previously participate in the market. Legal instabilities, carrying cost attributable to fluctuating prices and irreversibility represent possible factors that limit the spread of BTC.

Understanding that non-regulated digital currencies are indeed financial assets mitigates the potential threat to national legal tender currencies. The more restrictive laws therefore tend to be stricter for financial managers and traders. The comparison of BTC with the payment card market based on the 2SM structure showed that the use of BTC for obtaining credit has no similarity to the traditional model. However, when considering the debit system, the structure was adequate. This result indicates that BTC is considered an asset when incorporated into the financial system. In this sense, the processes of BTC price formation and the volatility associated with it were estimated using models that considered conditional heteroskedasticity. Each was estimated for the sample and subsample, as described above. The estimation of quasi-correlations between volatilities made it possible to obtain dynamic conditional volatility parameters of the DCC MGARCH (1,1) model. Analysis of the results shows that with the passage of time, the market has absorbed the BTC innovation.

Although attacks on financial institutions that transact BTCs are also elements that significantly affect price volatility, they adversely affect price. Because of the relationship between peaks in the price of BTC and epicenters of political and economic crises, it is suggested that future studies aiming to measure the volatility of BTC prices should incorporate variables into a multivariate approach capable of indicating the imminence of crises. Despite the impossibility of predicting economic cycles deterministically, analysis of macroeconomic fundamentals can provide BTC speculators with information useful for changing their strategies.

REFERENCES

- ALBUQUERQUE, Bruno Saboia de; CALLADO, Marcelo de Castro. **Understanding Bitcoins: facts and questions**. Revista Brasileira de Economia, v. 69, n. 1, p. 3-16, 2015.
- BANCO CENTRAL DO BRASIL; SECRETARIA DE ACOMPANHAMENTO ECONÔMICO; SECRETARIA DE DIREITO ECONÔMICO. **Relatório sobre a indústria de cartões de pagamentos**. Brasília, 2010.
- BAUWENS, Luc; LAURENT, Sébastien; ROMBOUTS, Jeroen VK. **Multivariate GARCH models: a survey**. Journal of applied econometrics, v. 21, n. 1, p. 79-109, 2006.
- BOLLERSLEV, T. **Generalized Autoregressive Conditional Heteroskedasticity**. Journal of Econometrics, 31, pp. 307-327, 1986.
- _____. **Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model**. The Review of Economics and Statistics, p. 498-505, 1990.
- BOLLERSLEV, Tim; ENGLE, Robert F.; WOOLDRIDGE, Jeffrey M. **A capital asset pricing model with time-varying covariances**. The Journal of Political Economy, p. 116-131, 1988.
- BOLLERSLEV, Tim; ENGLE, Robert F.; NELSON, Daniel B. **ARCH models**. Handbook of econometrics. V. 4, p. 2959-3038, 1994.
- ENGLE, Robert F. **Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation**. Econometrica: Journal of the Econometric Society, p. 987-1007, 1982.
- ENGLE, Robert. **Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models**. Journal of Business & Economic Statistics, v. 20, n. 3, p. 339-350, 2002.

EVANS, David S. **The antitrust economics of two-sided markets**. Available at SSRN 332022, 2002.

EVANS, David S. **Some empirical aspects of multi-sided platform industries**. Review of Network Economics, v. 2, n. 3, 2003.

FAGUNDES, J.; FERRÉS, J.; SAITO, K. **Indústria de cartões de crédito, regulação e concorrência**. Revista do IBRAC, v. 15, n. 4, 2009.

FRASCAROLI, Bruno Ferreira. **Uma investigação sobre moeda, meios de pagamentos e crédito no Brasil utilizando simulações nas taxas de juros: o que podemos dizer sobre as recentes contribuições em economia monetária?** 2010. 162 f. Tese (Doutorado em Economia) UFPE.

FREITAS, P. S. **Mercado de cartões de crédito no Brasil: problemas de regulação e oportunidades de aperfeiçoamento da legislação**. Brasília, DF: Consultoria Legislativa do Senado Federal, 2007. Texto para discussão n. 37, Brasília, DF.

MAY, Timothy C. **Crypto anarchy and virtual communities**. 1994.

MEIKLEJOHN, Sarah et al. **A fistful of bitcoins: characterizing payments among men with no names**. In: Proceedings of the 2013 conference on Internet measurement conference. ACM, 2013. p. 127-140.

NELSON, Daniel B. **Conditional heteroskedasticity in asset returns: A new approach**. Econometrica: Journal of the Econometric Society, p. 347-370, 1991.

OLIVEIRA, M. F.; TORKOMIAN, Ana Lucia Vitale. **Aprendizagem e inovação em pequenas empresas inseridas em aglomerações produtivas**. Congrès de L'Institut Franco-Brésilien d'Administration des Entreprises, 5^o., 2009, Grenoble.

PAAR, Christof; PELZL, Jan. **Understanding cryptography: a textbook for students and practitioners**. Springer Verlag, 2010.

PAVEL, Ciaian; D'ARTIS, Kancs; MIROSLAVA, Rajcaniova. **The Economics of BitCoin Price Formation**. EERI Research Paper Series, 2014.

ROCHET, Jean-Charles; TIROLE, Jean. **Cooperation among competitors: Some economics of payment card associations**. Rand Journal of economics, p. 549-570, 2002.

ROCHET, Jean-Charles; TIROLE, Jean. **Platform competition in two-sided markets**. Journal of the European Economic Association, p. 990-1029, 2003.

ROCHET, Jean-Charles; TIROLE, Jean. **Two-sided markets: a progress report**. The RAND Journal of Economics, v. 37, n. 3, p. 645-667, 2006.

SATOSHI, Nakamoto. **Bitcoin: a peer-to-peer electronic cash system**. (2008). Disponível em: <<https://bitcoin.org/bitcoin.pdf>>. Acesso em 16 de Agosto de 2015.

SCHILLER, R. J.. **Radical financial innovation**. In: Entrepreneurship, Innovation, and the Growth Mechanism of the Free-Enterprise Economies. Princeton University Press, 2007, v. p. 306-323.

SCHMALENSEE, Richard; EVANS, David S. **The economics of interchange fees and their regulation: An overview**. 2005.

STRÅLE, J. N; TJERNSTRÖM, M. **The price volatility of Bitcoin: A search for the drivers affecting the price of this digital currency**. 2014. 79 f. Dissertação. (Dissertação) Umeå Universitet. 28 de Outubro de 2014.

TAYLOR, Michael Bedford. **Bitcoin and the age of bespoke Silicon**. In: Proceedings of the 2013 International Conference on Compilers, Architectures and Synthesis for Embedded Systems. IEEE Press, 2013. p. 16.

TSAY, R. S. **Analysis of financial time series**. 2 ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2005.

WONGLIMPIYARAT, Jarunee. **The dynamics of financial innovation system**. The Journal of High Technology Management Research, v. 22, n. 1, p. 36-46, 2011.

WORLD ECONOMIC FORUM and DELOITTE. **The Future of financial services: how disruptive innovations are reshaping the way financial services are structured, provisioned and consumed**. 2015.

IMPACTOS DO PROGRAMA BOLSA FAMÍLIA NO MERCADO DE TRABALHO E NA RENDA DOS TRABALHADORES RURAIS

Edward Martins Costa¹

Rayssa Alexandre Costa²

Francisca Zilania Mariano³

Daniella Medeiros Cavalcante⁴

Grupo 2. Economia Social e do Trabalho

Resumo: O presente estudo avalia os efeitos do Programa Bolsa Família no mercado de trabalho das famílias pobres que residem em áreas rurais do Brasil. Para a análise, foi utilizado o Censo Demográfico (2010) e duas metodologias: *Propensity Score Matching* (PSM) e Efeito Quantílico de Tratamento (EQT). Observou-se como resultados do efeito médio, que no Brasil e regiões, as horas trabalhadas dos beneficiários são menores que a dos não beneficiários, assim como, a renda do trabalho. Quanto aos efeitos na distribuição, para a maioria dos *quantis* é nulo, entretanto, em alguns *quantis*, os beneficiários possuem menos horas trabalhadas que os não beneficiários. Os resultados, na média, indicaram um possível “efeito desincentivo” do programa sobre as horas trabalhadas, ocasionando uma redução na renda das famílias proveniente do trabalho. Na análise por *quantis* não se pode afirmar o mesmo, visto que os resultados dos efeitos nas horas trabalhadas foram distintos.

Palavras-chave: Programa Bolsa Família, *Propensity Score Matching*, Efeito Quantílico de Tratamento, Mercado de Trabalho.

Abstract: This study evaluates the effects of the Bolsa Família Program in the labor market of poor families living in rural areas of Brazil. For the analysis, we used the Census (2010) and two methodologies: Propensity Score Matching (PSM) and Quantílico treatment effect (EQT). It was observed as the average effect results in Brazil and regions, the working hours of the beneficiaries is lower than that of non-beneficiaries, as well as labor income. The effects on distribution, for most quantile the effect is null, however, in some quantile, beneficiaries have fewer working hours than non-beneficiaries. The results, on average, indicated a possible "deterrent effect" of the program on the hours worked, causing a decrease in household income from work. In the analysis by quantile can not say the same, as the results of the effects on hours worked were different.

Key words: Bolsa Família Program; Propensity Score Matching; Quantílico Treatment Effect; Label Market

CLASSIFICAÇÃO JEL: C21, C35, I30, R23

¹Doutor em Economia (PIMES/UFPE). Professor do Mestrado Acadêmico em Economia Rural (MAER/UFC). E-mail: edwardcosta@ufc.br.

²Mestre em Economia Rural (MAER/UFC). Doutoranda em Economia Aplicada (ESALQ/USP). E-mail: rayssacosta_@hotmail.com.

³Professora UFC/Campus de Sobral. Doutoranda em Economia, CAEN/UFC. E-mail: zilania@ufc.br.

⁴Mestre em Economia (PPGE/UFRN). Doutoranda em Economia (PIMES/UFPE). E-mail: dayukarini@gmail.com.

1. Introdução

O meio rural vem passando por diversas transformações nas últimas décadas, dentre elas, está a adoção do modelo de modernização do setor agrícola, ocorrida em meados da década de 1960, com a introdução de meios de produção mais avançados e de políticas públicas destinadas a este setor. Consequentemente esse período marcou o início de um novo modelo econômico no país por meio da mudança do chamado modelo de substituição de importações pela modernização do setor agrário e formação do Complexo Agroindustrial (TEIXEIRA, 2005).

De acordo com Ferreira *et al* (2006), embora tenha havido uma mudança produtiva na agropecuária brasileira, essa mudança se deu de forma desigual entre as regiões do país. Goodman, Sorj e Wilkison (1985 *apud* NASCIMENTO *et al*, 2008), destacam que os médios e grandes produtores foram os mais beneficiados, principalmente aqueles que se direcionavam para a produção de produtos de exportação e/ou ligados às agroindústrias, em especial os do Centro-Sul do país. Conforme Teixeira (2005), os produtores das pequenas propriedades não foram incluídos nesse processo de modernização, devido aos altos custos de produção. Assim, alguns pequenos produtores (principalmente aqueles que praticavam agricultura de subsistência) abandonaram as atividades agrícolas, deixando o campo e migrando para os centros urbanos.

Já na década de 1990, a abertura comercial brasileira foi implementada pelo governo Collor, na qual foi eliminada a maior parte das barreiras não-tarifárias e reduziram gradativamente o nível e o grau de proteção da indústria local (AVERBUG, 1999). Com isso, a entrada de produtos agrícolas similares aos nacionais foi permitida, aumentando a concorrência no mercado interno (CARVALHO; MARINHO, 2003). Aliada a abertura comercial, outras mudanças sofridas no período foram responsáveis pela diminuição da oferta de empregos no setor agrícola como a redução de incentivos do governo, a sobrevalorização do câmbio e a queda do preço dos produtos agrícolas (SANTOS *et al*, 2010).

Todos esses desdobramentos ampliaram ainda mais as disparidades de renda no meio rural. Incapazes de competir com a grande indústria rural, os pequenos produtores que escolheram continuar no campo, passaram a buscar outras formas de sobrevivência além da produção para o autoconsumo, assim, conforme Santos *et al* (2010), atividades não agrícolas, como artesanato, cultivo de flores, ecoturismo, prestação de serviços e comércio, viraram fontes de renda, viabilizando a permanência nas áreas rurais.

Diante disto, o meio rural apresenta uma grande distorção em relação à área urbana quanto a alguns indicadores, como menor rendimento médio familiar e maior percentual de famílias pobres e extremamente pobres (IBGE, 2015). Na tentativa de reverter ou amenizar esse quadro, algumas estratégias vêm sendo empregadas na esperança de fortalecer e promover o segmento rural, como políticas de crédito e programas de transferências de renda, em especial, o Programa Bolsa Família (PBF).

O PBF é um programa de transferência de renda condicionada que tem como objetivo atenuar a situação de pobreza das famílias, no curto prazo, com a imposição de condicionalidades nas áreas da educação, saúde e assistência social.

As maiores críticas e questionamentos são a respeito de um possível efeito negativo que o programa possui sobre o mercado de trabalho, o chamado “efeito desincentivo” (COSTA *et al*, 2014) ao trabalho. Segundo Tavares (2008), este efeito, também conhecido como efeito-renda, trata da redução da oferta de trabalho dos membros adultos das famílias beneficiadas em vista das transferências monetárias recebidas por estes. Ou seja, existe um tipo de recompensa pela perda salarial, por meio da transferência de renda repassada às famílias, assim, a renda familiar não é comprometida.

Diante do exposto, surgem alguns questionamentos: o Programa Bolsa Família gera “efeito-desincentivo” ao trabalho nas famílias beneficiadas pelo programa situadas nas áreas

rurais? Caso provoque, esse impacto é observado em todos os estratos de renda e nas horas trabalhadas?

Dessa forma, este trabalho busca investigar o efeito do PBF no mercado de trabalho, por meio das horas trabalhadas e da renda, das famílias pobres situadas nas áreas rurais por meio de duas metodologias: *Propensity Score Matching (PSM)* e Efeito Quantílico de Tratamento (EQT) proposto por Firpo (2007). Diante disso, a contribuição deste trabalho para a literatura se apresenta, principalmente, por mensurar o efeito do programa, especificamente para as áreas rurais das regiões brasileiras, na média e em toda a distribuição, permitindo verificar os estratos de renda e as horas de trabalho semanais, nos quais o efeito é mais expressivo.

Neste sentido, este trabalho está estruturado em cinco seções, incluindo esta introdução. Na seção dois, será feita uma discussão teórica sobre o Programa Bolsa Família, os efeitos do PBF no mercado de trabalho, assim como alguns estudos empíricos que foram feitos. A metodologia será apresentada na terceira seção. Os resultados podem ser observados na seção quatro. E, por último, as considerações finais do trabalho serão apresentadas na seção cinco.

2 REVISÃO DE LITERATURA

2.1 Programa Bolsa Família - PBF

O Programa Bolsa Família⁵ (PBF) é um programa de transferência direta de renda condicionada e foi criado pelo Governo Federal com o objetivo de, no curto prazo, aliviar a situação de pobreza e promover segurança alimentar e nutricional das famílias que se encontram em posição de vulnerabilidade social em todo o país. Segundo Soares e Sátyro (2009), o PBF resultou da unificação⁶ de outros programas de transferências de renda (Bolsa Escola, Bolsa Alimentação, Auxílio Gás e Cartão Alimentação), que à época se encontravam em situação de caos, onde cada programa federal tinha sua agência executora, além disso, a coordenação entre elas era mínima. Os sistemas de informação dos programas eram distintos e não havia comunicação entre eles, como consequência muitas famílias poderiam receber benefícios de todos os programas, enquanto outras que se encontravam nas mesmas condições sociais poderiam ser negligenciadas, podendo até mesmo, não receber qualquer benefício.

O Bolsa Família, desde a sua criação, conta com o CadÚnico como uma fonte de informações (registradas pelos municípios) para o programa e para qualquer outro programa social focalizado, que facilita a identificação das famílias de baixa renda existentes no Brasil, como também, a seleção das famílias que poderão receber o benefício (BRASIL, 2015). O critério de renda utilizado pelo Governo Federal é dado mediante uma linha de pobreza e extrema pobreza⁷, sendo classificadas como pobres aquelas famílias que recebem entre R\$ 77,01 e R\$ 154,00 *per capita* e famílias extremamente pobres aquelas que possuem renda inferior a R\$77,00 *per capita*. Se as famílias estão inseridas nessas faixas de rendas, então elas podem ser consideradas como elegíveis e dependendo da quantidade de filhos e de suas respectivas faixas etárias, os benefícios destas podem variar. Vale ressaltar, que para famílias extremamente pobres é concedido um benefício básico, no valor de R\$77,00, que não exige condicionalidades. Portanto, existem duas modalidades de benefícios: **Benefício Básico:** Concedido às famílias extremamente pobres; **Benefício Variável:** Concedido às famílias pobres, que tenham crianças de 0 a 15 anos, gestantes e nutrízes, no valor de R\$35,00 e com

⁵ Instituído pela Medida Provisória nº 132, de 20 de outubro 2003. Oficializado pela Lei nº 10.836, de 09 de janeiro de 2004, e Decreto nº 5.209, de 17 de setembro de 2004.

⁶ A partir de 2006, o Programa de Erradicação do Trabalho Infantil (PETI) também passou a compor o PBF.

⁷ Estes valores correspondem à linha de pobreza do ano de 2105. Para o ano de análise deste estudo, 2010, os valores correspondiam a: até R\$ 70,00 (famílias extremamente pobres) e entre R\$70,01 e R\$140 (famílias pobres).

um limite de 5 benefícios por família. Há, também, o benefício para adolescentes de 16 e 17 anos, no valor de R\$42,00 e limitado a 2 benefícios por família.

Embora existam incentivos governamentais para amenizar a situação de pobreza e risco das famílias mediante programas de transferências de renda, como o PBF, nem todas as famílias que se encontram nesta situação são contempladas com o benefício, visto que há uma limitação dos recursos governamentais.

Até 2006, o PBF foi avançando gradualmente, alcançando a meta de 11 milhões de famílias beneficiadas (SOARES; SÁTYRO, 2009). Conforme os dados do Sistema de Benefícios ao Cidadão – SIBEC (2015), o programa beneficiou em 2015 aproximadamente 14 milhões de famílias em todo território nacional, sendo um dos programas sociais que maior contempla e colabora com a redução da miséria. Segundo o Ministério de Desenvolvimento Social e Combate à Fome – MDS (2015), o PBF visa assegurar o acesso a direitos sociais básicos, nas áreas de educação, saúde e assistência social. Para que as famílias permaneçam no programa recebendo o benefício, é necessário que elas cumpram com algumas condicionalidades exigidas pelo Governo Federal nestas áreas supracitadas. Ressalta-se que o MDS, juntamente com os Ministérios da Educação e da Saúde, são os órgãos responsáveis por fazer o monitoramento das condicionalidades do programa, cabendo também aos municípios fazer o acompanhamento intersetorial entre as áreas (saúde, educação e assistência social).

Na área da educação, a contrapartida exigida das famílias é que todas as crianças entre 6 e 15 anos estejam devidamente matriculadas e que tenham frequência mensal mínima de 85%, enquanto os jovens de 16 e 17 anos devem ter frequência mínima de 75%. Tratando-se da saúde, as famílias devem acompanhar o cartão de vacinação de crianças menores de 7 anos. Já as mulheres, de 14 a 44 anos, gestantes ou nutrizes, assumem o compromisso de fazer o pré-natal e o acompanhamento de sua saúde e do bebê. Quanto à assistência social, é exigido que todas as crianças menores de 15 anos que se encontram ou encontravam-se em situação de risco de trabalho infantil participem do Programa de Erradicação do Trabalho Infantil – PETI e obtenham frequência escolar mínima de 85%.

O não cumprimento das contrapartidas impostas às famílias não necessariamente implica na perda do benefício de imediato. Dependendo do motivo que ocasionou o não cumprimento das condicionalidades, serão atribuídas penalidades às mesmas e, caso persista o descumprimento por parte das famílias, o cancelamento do benefício pode ser feito. Soares e Sátyro (2009) ressaltam que, para as famílias que se encontram em situação de extrema vulnerabilidade, é mais difícil que essas contrapartidas sejam obedecidas, visto que algumas famílias vivem longe de escolas e postos de saúde. Portanto, é de suma importância que o Estado cumpra com seu dever constitucional de criar condições para que estas famílias possam desempenhar com êxito o que lhes foi imposto e, assim, melhorar as suas condições de vida.

2.2 Efeitos dos Programas de Transferências de Renda sobre a Oferta de Trabalho

Um dos temas muito debatido, envolvendo transferências de renda, é o impacto que esses tipos de programas sociais tem sobre o mercado de trabalho das famílias pobres beneficiadas. Levy (2010) argumentou a respeito de um “ciclo vicioso” de programas sociais não contributivos, que acabam incentivando as famílias a procurarem empregos de menor qualidade ou de baixa produtividade, ou seja, os trabalhadores acabam ingressando no setor informal, para assim evitar contribuições obrigatórias à previdência social. Entretanto, segundo a Cepal (2006 *apud* CECCHINI 2013), os programas de proteção social são importantes fontes de investimento em capital humano e o que gera e mantém o setor informal na economia é a heterogeneidade da estrutura produtiva, no qual os setores modernos, de maior produtividade, tem pouca capacidade de absorver a mão-de-obra. Portanto, muitas vezes o setor informal é a única possibilidade de inserção laboral dos trabalhadores.

Cecchini (2013) mostra que existe um determinado elo positivo entre a proteção social não contributiva, o emprego e os princípios do trabalho decente. Para a autora, os programas de transferências condicionadas acabam proporcionando uma maior liquidez às famílias beneficiadas, fazendo com que estas sejam capazes de tomar melhores decisões com relação a condições dignas de trabalho e empregabilidade.

As principais críticas feitas, especialmente quanto ao Programa Bolsa Família, são acerca de um possível impacto negativo sobre os membros adultos das famílias que possuem idade para trabalhar, devido a uma determinada acomodação desses indivíduos, diminuindo assim a oferta de trabalho destes, em virtude do recebimento de benefícios básicos (benefícios que não exigem uma contrapartida das famílias). Ou seja, esse tipo de transferência gera um “efeito-desincentivo” ao trabalho nas famílias (COSTA *et al*, 2014). Brito (2011) ressalta que tal efeito produz uma certa dependência das famílias beneficiadas em relação ao Estado.

Cecchini (2013) afirma que os programas de transferência renda condicionada não são suficientes para arcar com o déficit da renda das famílias beneficiadas, devido ao montante repassado ser um valor baixo. Para Tavares (2008), existem dois tipos de efeito que podem ser observados em programas de transferências de renda como o PBF: Efeito-renda: Neste efeito, tem-se uma redução na oferta de trabalho, devido a um aumento da renda da família diante do recebimento do benefício do programa. Dessa forma, a renda da família não é comprometida, ou seja, existe uma recompensa pela redução da oferta de trabalho proveniente da transferência de renda do PBF. Este tipo de efeito também é conhecido como “efeito-desincentivo”. Efeito-substituição: Este efeito decorre da imposição das condicionalidades. Caso a renda familiar seja oriunda de todos os membros da família e, visto que, o programa impõe uma contrapartida quanto à frequência escolar das crianças e jovens, outro membro familiar deverá aumentar a oferta de trabalho para compensar a perda da renda devido à redução do trabalho dessas crianças e jovens. Dado que os dois efeitos atuam em direções opostas, tem-se um equilíbrio quando o efeito-renda é igual ao efeito-substituição. Caso o efeito-renda sobreponha o efeito-substituição, tem-se um impacto negativo do programa, e caso contrário, um impacto positivo.

Alguns trabalhos empíricos foram realizados com a intenção de mostrar possíveis impactos dos Programas de Transferência de Renda no mercado de trabalho. A seguir, serão descritos alguns desses estudos. Parker e Skoufias (2000) avaliaram o impacto do Programa de transferência de renda Progres⁸, no México, na participação no mercado de trabalho, no lazer e alocação do tempo dos homens, mulheres e crianças, considerando a metodologia de diferenças em diferenças probit e os dados do censo ENCASEH⁹ de 1997, 1998 e 1999. No caso da participação no mercado de trabalho, os resultados mostraram que houveram reduções significativas na participação da força de trabalho infantil em atividades assalariadas e não assalariadas tanto para meninos quanto para meninas. Quanto aos adultos, homens e mulheres, não foram constatados efeitos de desincentivo ao trabalho. Para os resultados da alocação de tempo, foram observados que as crianças aumentaram significativamente a participação nas atividades escolares. Com relação aos adultos, não há impacto significativo do programa no tempo de lazer.

Cardoso e Souza (2004) estimaram o impacto sobre a frequência escolar e o trabalho infantil de programas de transferências de renda, como o programa de renda mínima e o bolsa escola, para famílias pobres no Brasil. Por meio dos dados do Censo (2000) e a metodologia

⁸ O Programa de Educación Salud y Alimentación (Progres) iniciou em 1997, posteriormente, em 2002, recebeu o nome de Oportunidades e em 2014 passou a se chamar PROSPERA

⁹ O ENCUESTA DE CARACTERÍSTICAS SOCIOECONÓMICAS DE LOS HOGARES (ENCASEH) é um questionário aplicado para saber quais famílias podem ser atendidas pelo PROPSRA, antigo PROGRESA, no qual é feita a coleta de dados sobre saúde, educação, alimentação, trabalho, acesso aos serviços básicos, características da habitação, posse de propriedade, o tipo de localidade, entre outro.

do PSM, encontraram-se resultados iguais tanto para meninos quanto para meninas com relação ao aumento da frequência escolar. Tratando-se da probabilidade de trabalhar, observou-se uma diminuição significativa para as meninas. Já para os meninos, os resultados não foram estatisticamente significantes.

O trabalho feito pelo Cedeplar (2006) avaliou o impacto do PBF sobre a oferta de trabalho, com base em uma pesquisa de campo. Os resultados indicaram um efeito positivo do programa, mostrando que os adultos membros das famílias beneficiadas tem uma taxa de participação maior do que os adultos residentes em domicílios não beneficiários. Esse efeito ainda é maior com relação às mulheres, quando comparado aos homens.

Ferro e Nicollela (2007) analisaram o efeito das transferências de renda condicionadas quanto à taxa de participação e as horas trabalhadas dos adultos nas áreas urbanas e rurais, com base nos dados da PNAD do ano de 2003. Os efeitos encontrados foram positivos e significativos apenas para homens e mulheres residentes em áreas urbanas e negativo para mulheres residentes em áreas rurais. Teixeira (2008) estima o efeito do PBF na oferta de trabalho em atividades formais e informais de homens e mulheres beneficiários do programa, com dados da PNAD 2006. O cálculo é feito por meio do efeito médio do tratamento (*Average Treatment Effect on the Treated – ATT*) e a análise gráfica do efeito local do programa (*Average Local Effect – ALE*). Os resultados apontam que o PBF tem efeito negativo na oferta de trabalho das mulheres. Por sua vez, na oferta de trabalho dos homens, o efeito é negativo para os que trabalham por conta própria e nulo para os que trabalham no mercado formal.

Tavares (2008) investigou o impacto do PBF sobre a oferta de trabalho das mães beneficiadas. Foram utilizados dados da Pesquisa Nacional por Amostra Domiciliar (PNAD), referentes ao ano de 2004, e o método do *Propensity score Matching (PSM)* para assim fazer um *confractual* e encontrar mães que não foram beneficiadas pelo programa comparáveis às mães que foram. Os resultados encontrados evidenciam um pequeno efeito-renda negativo, mas que não é suficiente para gerar um *efeito-desincentivo* ao trabalho. Entretanto, também foi observado um efeito positivo que se sobrepõe ao efeito negativo, a autora explica que isso pode decorrer devido a uma maior disponibilidade de tempo das mães para trabalhar, visto que as crianças e jovens tem um aumento da frequência escolar e assim reduzem sua oferta de trabalho.

Mattos *et al* (2010) fizeram a análise da influência de dois programas sociais (Programa de Renda Mínima e Bolsa Escola) quanto à oferta de trabalho dos beneficiários. Os autores utilizaram a metodologia de seleção de Heckman (1979) e o PSM com dados da Pesquisa de Orçamentos Familiares (POF) no período de 2002 e 2003. Como resultado, não ficou evidenciado, entre os beneficiários, diferença com relação entre a probabilidade de estar empregado e o número de horas trabalhadas no ano. E ainda, houve um pequeno efeito-negativo na probabilidade de estar empregado das mulheres que recebiam benefícios do Programa Renda Mínima e um efeito-positivo para os homens que recebiam benefícios do Bolsa Escola.

Gonzalez-Rozada e Pinto (2011) investigaram o impacto do Programa Bono de Desarrollo Humano (BDH), do Equador, sobre o mercado de trabalho dos beneficiários. Utilizando dados do ENEMDU¹⁰ para os anos de 2004 a 2010 e a metodologia de regressão descontínua, os autores encontraram que para as mães beneficiárias e os trabalhadores que pertencem a famílias que recebem os benefícios, obtiveram maior duração do desemprego do que os não beneficiários, além disso, o programa não obteve efeitos de distorção sobre a probabilidade de existência de trabalho informal para as mães e os trabalhadores que vivem em famílias com benefícios do BDH.

¹⁰ Pesquisa Nacional de Emprego, Desemprego e Subemprego realizada pelo Instituto Nacional de Estatística e Censos (INEC) do Equador.

Alzúa *et al* (2012) avaliaram o efeito de três programas implementados em áreas rurais (Progres – México, Red de Protección Social – Nicarágua, Programa de Asignación Familiar – Honduras) sobre a oferta de trabalho de adultos beneficiários. Os dados utilizados foram provenientes de pesquisas longitudinais *ad hoc* realizadas a fim de avaliar cada intervenção específica, ademais, as três fontes de dados foram harmonizadas com base em um conjunto comum de critérios a fim de alcançar o máximo de compatibilidade com a metodologia descrita no CEDLAS¹¹ (2012). Por meio da técnica de estimação de diferenças em diferenças, os resultados da oferta de trabalho para os adultos que participaram dos programas foram em sua maioria negativos, entretanto foram efeitos pequenos e alguns foram estatisticamente não significantes. Também foram observadas algumas particularidades, como por exemplo no caso do PROGRESA em que há um pequeno efeito positivo sobre o número de horas trabalhadas dos beneficiários do sexo feminino.

Costa *et al* (2014) fizeram uma análise do impacto do PBF na renda e nas horas de trabalho das famílias pobres em todas as regiões brasileiras. Para a análise, foram utilizados dados do Censo (2010) do IBGE e o método do Efeito Quantílico do Tratamento (EQT), em especial o estimador proposto por Firpo (2007). Os resultados mostraram que em grupos específicos de famílias beneficiadas houveram menos horas trabalhadas e uma menor renda do trabalho quando comparado às famílias não beneficiadas. No tocante às regiões, os resultados foram distintos, tendo as regiões Nordeste e Norte efeitos negativos no mercado de trabalho, enquanto os resultados para as demais regiões foram mais homogêneos. Também foi possível observar efeitos para o Brasil urbano e Brasil rural, no qual os impactos foram diferentes, com efeitos neutros, positivos e negativos para os dois casos, entretanto o efeito desincentivo ao trabalho foi mais forte no Brasil rural do que no urbano, especialmente nos *quantis* mais elevados. Os autores indicam que parte dessas diferenças observadas pode ter ligação com o dinamismo econômico específico de cada região.

3 METODOLOGIA E BASE DE DADOS

3.1 Base de Dados

Os dados utilizados neste trabalho foram retirados do Censo Demográfico, do Instituto Brasileiro de Geografia e Estatística (IBGE), para o ano de 2010¹². Foram escolhidas variáveis que estavam presentes simultaneamente no Censo Demográfico e no CadÚnico. Para obter os efeitos do Programa Bolsa Família no mercado de trabalho, foram utilizadas duas metodologias que avaliam o efeito em situações distintas. A primeira metodologia, *Propensity Score Matching*, analisa os efeitos do programa na média e a segunda, Efeito Quantílico de Tratamento de Firpo (2007), verifica os efeitos ao longo da distribuição. A seguir, serão abordadas as duas metodologias.

3.2 Modelos Econométricos

3.2.1. *Propensity Score Matching (PSM)*

Na avaliação de políticas públicas, o ideal para identificar o efeito de uma política seria comparar um mesmo grupo em dois contextos: participando e não participando do programa. Entretanto, isso não é possível e o que pode ser feito é criar um contrafactual, que seria um grupo estatisticamente idêntico ao grupo que é beneficiado pelo programa, diferindo apenas do fato de não receber o benefício.

O *Propensity Score Matching* (PSM) é um dos métodos existentes para este tipo de análise, no qual consiste na seleção de um grupo de controle, com base em um modelo de probabilidade (Logit/Probit) de participar do tratamento, comparáveis com um grupo de

¹¹ Centro de Estudos Distributivos, Laborais e Sociais da Faculdade de Ciências Econômicas da Universidad Nacional de La Plata

¹² Visto que foi o último censo divulgado.

tratamento por meio do pareamento (*matching*) dos grupos, utilizando características¹³ observáveis. A partir disso, pode-se comparar os resultados dos tratados e não tratados com escores de propensão semelhantes para obter o efeito do programa. O escore de propensão, desenvolvido por Rosenbaum e Rubin (1983), é definido como a probabilidade condicional de receber um tratamento, dado por características do pré-tratamento observáveis, isto é:

$$p(X) = \Pr(T = 1|X) = E(T|X) \quad (1)$$

Onde T indica a posição do tratamento para a família (1 – recebe o benefício do programa, 0 – não recebe) e X é o vetor de características. Assim, o efeito médio do tratamento sobre os tratados (ATT), é dado por:

$$\tau = E_{p(X)|T=1}\{E[Y_{1i}|T_i = 1, p(X_i)] - E[Y_{0i}|T_i = 0, p(X_i)]|T_i = 1\} \quad (2)$$

Para que se possa calcular o ATT, é necessário que sejam atendidas duas hipóteses. A primeira hipótese diz a respeito à independência condicional, que implica que os fatores não observados não afetam a participação. No caso do Bolsa Família, os beneficiários são determinados mediante sua inscrição no CadÚnico, posteriormente é realizado um sorteio aleatório para a concessão do benefício. **Hipótese 1:** $Y_i(1), Y_i(0) \perp T_i | X_i$

A segunda hipótese trata do suporte comum, que conforme Heckman, LaLonde e Smith (1999), este pressuposto garante que as observações de tratamento têm observações de comparação "nas proximidades" na distribuição de escore de propensão. Para atender essa hipótese, as variáveis escolhidas para o *matching* foram aquelas com mais similaridades entre os beneficiários e não beneficiários. **Hipótese 2:** Para algum $c > 0$, $c < p(x) < 1 - c$.

De acordo com Rosenbaum e Rubin (1983), essas duas hipóteses combinadas são conhecidas como ignorabilidade forte. Além disso, ressalta-se que a estimação do escore de propensão não é suficiente para estimar o ATT, sendo necessário aliá-la ao *matching*. Com relação a isso, diferentes métodos podem ser utilizados para fazer o pareamento, dentre eles, destacam-se os que seguem: Pareamento por vizinho mais próximo (*Nearest Neighbor Matching*), Pareamento por alcance (*Radius Matching*), Pareamento Kernel (*Kernel Matching*).

3.2.2 Efeito Quantílico de Tratamento – Modelo de Firpo (2007)

O modelo de efeito quantílico de tratamento exógeno e não condicional proposto por Firpo (2007) será utilizado para obter o efeito do PBF nas áreas rurais. Isto porque, uma família pode ou não pertencer ao programa devido a fatores aleatórios, dada uma restrição de renda e características observáveis, portanto, o tratamento utilizado deve ser exógeno. E não condicional deve-se ao fato de muitas famílias serem elegíveis ao programa, mas que não são selecionadas devido aos recursos direcionados ao programa serem limitados, dessa forma, a participação ou não condicional das famílias é feita de forma aleatória.

Neste estudo, o tratamento diz respeito a receber ou não o PBF, portanto os beneficiários do PBF fazem parte do grupo de tratamento e os não beneficiários, do grupo de controle. Visto isso, o modelo de EQT de Firpo (2007) é definido a seguir.

Seja T a variável que indica o tratamento para um indivíduo i , que assume valores de 0 e 1. Se $T_i=1$ o indivíduo assume a forma de tratado e se $T_i=0$, de não tratado. Seja Y_i a variável de resposta desse evento, ou seja, os resultados potenciais de receber ou não o tratamento, dessa forma, se $Y_i(1)$ o indivíduo recebe o tratamento e se $Y_i(0)$, não recebe o tratamento. Dessa forma, define-se o resultado observado como:

$$Y_i = Y_i(1) - T_i + Y_i(0) - (1 - T_i) \quad (3)$$

Diante disto, definindo τ como um número real em $[0,1]$ o efeito quantílico de tratamento (EQT) é dado por:

¹³ O grupo de covariáveis utilizadas para a estimação do PSM serão descritas na subseção 3.5.

$$\Delta_{\tau} = q_{1,\tau} - q_{0,\tau}, \text{ onde } q_{j,\tau} \equiv \inf_q \Pr[Y(j) \leq q_{j,\tau}] \geq \tau, j = 0,1 \quad (4)$$

A expressão acima mostra que o EQT é a diferença entre os valores do *quantil* τ das funções de distribuição cumulativas da variável resposta referentes ao grupo de tratamento e o de controle, para um dado *quantil*. Para estimar o EQT é necessário seguir duas etapas: em primeiro lugar, estima-se o escore de propensão e posteriormente calcula-se a diferença entre tratados e não tratados. Os estimadores do escore de propensão são definidos como a probabilidade condicional de receber um tratamento, dado por características de pré-tratamento observáveis:

$$p(X) = \Pr(T = 1|X = x) \quad (5)$$

A utilização do escore de propensão requer que sejam adotadas três hipóteses: independência condicional, suporte comum e *quantis* bem definidos e únicos. A hipótese de independência condicional implica que os valores de Y_i independem da aplicação do tratamento, condicionada a um determinado conjunto de variáveis X observáveis: **Hipótese 1:** $Y_i(1), Y_i(0) \perp T_i | X_i$. Já a hipótese de suporte comum garante que as observações de tratamento têm observações de comparação "nas proximidades" na distribuição de escore de propensão (Heckman, LaLonde, e Smith, 1999), ou em outros termos: **Hipótese 2:** Para algum $c > 0$, $c < p(x) < 1 - c$. **Hipótese 3:** Para $j=0,1$, $Y(j)$ é uma variável aleatória contínua com suporte em $\mathbb{R}$, no qual há um conjunto não vazio de Y_1 e Y_1 , de tal forma que $Y_i = \{\tau \in (0,1); \Pr[Y(j) \leq q_{j,\tau} - c] < \Pr[Y(j) \leq q_{j,\tau} + c], \forall c \in \mathbb{R}, c > 0\}$. Essa hipótese que mostra que os *quantis* são bem definidos e únicos.

Cabe mencionar que diferentemente da metodologia exposta na subseção anterior, a metodologia de Firpo (2007) não faz o pareamento (*matching*) dos escores de propensão, e a estimação do escore de propensão é feita por meio de um modelo logit global. Dessa forma, o estimador de EQT é dado por $\hat{\Delta}_t = \hat{q}_{1,\tau} - \hat{q}_{0,\tau}$, onde:

$$\hat{q}_{j,\tau} \equiv \arg \min_q \sum_{i=1}^N \hat{w}_{j,\tau} \cdot \rho_{\tau}(Y_i - q) \quad (6)$$

De tal forma que os *quantis* podem ser estimados mediante a minimização de uma soma da função *check* $\rho_{\tau}(\cdot)$ assim como foi proposto por Koenker e Bassett (1978) *apud* Cavalcanti (2013). Os pesos de cada uma das observações, w_i , é dada pela seguinte expressão:

$$\hat{w}_{1,\tau} = \frac{T_i}{N \cdot \hat{p}(X_i)} \text{ e } \hat{w}_{0,\tau} = \frac{1-T_i}{N \cdot (1-\hat{p})(X_i)} \quad (7)$$

3.3 Grupos de Tratamento e Controle

A avaliação do impacto do Bolsa Família no mercado de trabalho será realizada por meio da comparação dos resultados de dois grupos: **Grupo de Tratamento:** Diz respeito às famílias que atendem aos critérios de elegibilidade e são beneficiárias do programa; **Grupo de Controle:** São aquelas famílias que atendem aos critérios de elegibilidade, mas que não são beneficiárias, pois não foram selecionadas por algum motivo.

3.4 Variável de Resultado

Segundo Costal *et al* (2014), os impactos dos Programa Bolsa Família no mercado de trabalho serão avaliados utilizando duas variáveis de resultado: **Horas trabalhadas do chefe familiar:** Neste caso, pretende-se analisar se o programa reduz ou não as horas trabalhadas dos chefes beneficiários, ou seja, se realmente há um efeito desincentivo ao trabalho devido ao repasse monetário. **Renda do trabalho per capita:** Para esta variável de resultado, a análise se dá com o intuito de verificar se a renda das famílias oriunda do trabalho está aumentando ou reduzindo por possíveis efeitos negativos, de desincentivo ao trabalho, do programa.

Dessa forma, serão estimados dois modelos para as duas metodologias descritas, o primeiro avaliará o impacto do programa nas horas trabalhadas do chefe de família e o segundo analisará o impacto na renda do trabalho *per capita*.

3.5 Covariadas

Para fazer uma comparação dos resultados dos grupos de tratamento e de controle, é necessário que os dois grupos sejam o mais homogêneo possível. Dessa forma, com o intuito de garantir que o modelo capte o efeito do programa, fez-se necessário controlar características observáveis do grupo familiar e do domicílio, sendo consideradas as variáveis que estavam presentes simultaneamente no Censo Demográfico e no CadÚnico. O grupo de Covariadas é dividido em quatro categorias, como mostra o quadro 1 abaixo.

Caraterísticas do chefe familiar	Idade
	Gênero (1=homem, 0=mulher)
	Cor (1=branco, 0=não branco)
	Cônjuge (1=casado(a), 0=não casado(a))
	Escolaridade (dummies)
	Trabalha (1=trabalha, 0=não trabalha)
	Carteira de trabalho assinada (1=possui, 0=não possui)
Composição familiar:	Nº de pessoas
	Nº de filhos (por faixa etária)
	Nº de filhos homens (1=filho homem, 0=caso contrário)
	Nº de filhos brancos (1=filho branco, 0= caso contrário)
Outras características dos componentes domiciliares	Deficiência auditiva (1=possui, 0=não possui)
	Deficiência física (1=possui, 0=não possui)
	Deficiência visual (1=possui, 0=não possui)
	Deficiência mental (1=possui, 0=não possui)
Outras características do município	Próprio (1=é, 0=caso contrário)
	Tipo de domicílio (1=casa/apartamento, 0=oca/barraca/cortiço)
	Domicílios de alvenaria (1=alvenaria, 0=caso contrário)
	Rede de água tratada (1=possui, 0=não possui)
	Eletricidade (1=possui, 0=não possui)
	Saneamento (1=possui, 0=não possui)
	Coleta de lixo (1=possui, 0=não possui)

Fonte: Elaborado pelos autores, com base na pesquisa.

Quadro1. Covariadas utilizadas para o matching

4 RESULTADOS

4.1 Estatísticas Descritivas

A distribuição dos domicílios rurais pobres elegíveis para o Programa Bolsa Família pode ser observada na Tabela1. Os dados amostrais denotam um total de 430.227 unidades domiciliares, e, quando ponderados pelo peso, os dados do Universo mostram que os números de domicílios correspondem a 2.742.705.

No Brasil, o percentual de famílias beneficiadas pelo PBF equivale a 59,60% do total de famílias elegíveis. Quando a análise é feita por regiões, percebe-se que o Nordeste é responsável pela maior quantidade de famílias beneficiadas, tanto em termos absolutos quanto relativos, com cerca de 68,68% da amostra de 261.500 domicílios e representando 72,03% do total dos beneficiários do país. Em seguida, destacam-se as regiões Norte e Sudeste, que participam com 12,98% e 8,93%, respectivamente, no total de beneficiários no Brasil, mas só contemplam 46,20% (Norte) e 46,94% (Sudeste) das famílias elegíveis em suas regiões. Já as regiões Sul e Centro-Oeste possuem participações menos expressivas representando 4,02% e 2,03% dos beneficiários no país e com menos de 40% de suas populações recebendo benefícios.

Tabela 1 – Estatísticas descritivas – Brasil e Regiões – 2010

Regiões	Amostra	Universo ¹	Beneficiários do PBF (%)	Participação dos Beneficiários do PBF no Total de Beneficiários do Brasil
Brasil	430.227	2.742.705	59,60%	100,00%
Nordeste	261.500	1.714.334	68,68%	72,03%
Norte	62.050	459.119	46,20%	12,98%
Sudeste	55.450	311.130	46,94%	8,93%
Sul	34.386	162.913	40,33%	4,02%
Centro-Oeste	16.841	95.210	34,93%	2,03%

Nota: (1) Os dados do Universo são obtidos por meio da ponderação com os pesos divulgados pelo Censo Demográfico. Fonte: Elaborado pelos autores, com base nos dados do Censo Demográfico (2010).

4.2 Resultados para o *Propensity Score Matching* (PSM)

4.2.1 Impacto do Programa Bolsa Família nas horas trabalhadas do chefe familiar

Para a análise do efeito do Bolsa Família nas horas trabalhadas pelo método PSM, estimou-se, primeiramente, um modelo de probabilidade logit da família participar ou não do PBF, de maneira que fosse possível encontrar um suporte comum entre os grupos, assim, todos aqueles que estiveram fora do suporte foram descartados. Os resultados deste modelo podem ser vistos na Tabela A1 do apêndice. Para o Brasil, as variáveis de deficiências auditiva e mental não foram estatisticamente significantes a 5%. As demais variáveis foram significantes a 5%, com exceção das variáveis de ensino primário e deficiência física que foram estatisticamente significantes a 10%. Além disso, quanto maior a idade do chefe, menor é a chance de participação no programa, o mesmo é observado para casos de ser chefe homem, chefe branco, possuir ensino ginásio, segundo grau, superior, pós graduação, carteira de trabalho assinada, ser agricultor e ter saneamento e tratamento de lixo.

O intervalo da região de suporte comum para o Brasil foi entre 0,07688124 e 0,99987008. Foram consideradas 103.619 famílias dentro deste intervalo, de uma total de 103.648. Essa região garante que as observações dos dois grupos, tratamento e controle, podem ser comparáveis quanto às características observáveis. O passo seguinte foi realizar o teste de balanceamento¹⁴ para o Brasil, o qual apresentou o método de *kernel*¹⁵ como o mais adequado para se fazer o pareamento das variáveis, visto que possui menos variáveis significantes a 5%, além disso, houve uma redução do Pseudo-R² de 0,126 para 0,002, isto implica que as variáveis explicativas perdem poder de explicação com relação as diferenças entre as famílias (Tabela A3 do apêndice).

Após o pareamento, foi possível calcular o efeito médio do tratamento sobre os tratados (ATT), que pode ser visto na tabela abaixo. Os resultados mostram que o impacto do PBF nas horas trabalhadas do chefe familiar, no Brasil, foi em média de -2,04, ou seja, os beneficiários do programa possuem menos horas trabalhadas do que os não beneficiários. O mesmo é observado para todas as regiões, sendo o maior impacto verificado nas regiões Sul, Centro-Oeste e Sudeste, no qual os beneficiários possuem 2,39, 2,04 e 1,73, respectivamente, a menos de horas trabalhadas comparados às famílias que não recebem o repasse monetário. Já nas regiões Nordeste e Norte, os impactos foram menos expressivos, no qual a média das horas trabalhadas entre os grupos foi de -1,20 e -1,15, respectivamente. Vale ressaltar que 2010 foi um ano de seca no Nordeste e que mais de 50% dos beneficiários são agricultores, assim pode ser que neste ano o repasse monetário do programa tenha ajudado a atender

¹⁴ O teste de balanceamento para as regiões mostrou que o método de Kernel é o mais adequado, com exceção da região Nordeste, no qual o método do vizinho mais próximo foi o que melhor se adequou. Os resultados para o teste de balanceamento podem ser observados no Apêndice.

¹⁵ O matching Kernel e Raio apresentaram o mesmo balanceamento, porém, o Kernel foi escolhido por não apresentar no pareamento famílias com escores de propensões diferentes. Desta forma, o grupo dos tratados e de controle serão mais homogêneos.

apenas as necessidades fundamentais dos beneficiários, sem que tenha havido grandes incentivos para uma maior redução das horas trabalhadas.

Embora os resultados das horas trabalhadas entre os beneficiários e o grupo de controle, tenham mostrado que a Região Sul é aquela que apresenta o maior “efeito desincentivo” médio, esta região é responsável pela maior quantidade de horas trabalhadas entre todas as regiões. Por outro lado, o Nordeste que apresentou o segundo menor “efeito desincentivo” é a região com a menor média de horas trabalhadas, além disso a região se destaca por ser responsável por aproximadamente 72% do total de beneficiários do Programa Bolsa Família no Brasil, como visto na tabela 1.

Tabela 2 - Impacto do PBF nas horas trabalhadas do chefe familiar

Regiões	Tratados	Controle	ATT	teste t
Brasil	31,31	33,35	-2,04	-15,04
Nordeste	29,28	30,48	-1,20	-6,83
Norte	31,22	32,36	-1,15	-3,52
Sudeste	33,95	35,68	-1,73	-4,71
Centro-Oeste	32,97	35,02	-2,04	-2,47
Sul	35,90	38,29	-2,39	-6,25

Fonte: Elaborado pelos autores, com base nas estimações.

4.2.2 Impacto do Programa Bolsa Família na renda do trabalho *per capita*

Seguindo os procedimentos realizados para a análise anterior, estimou-se um modelo logit, no qual os resultados estão expostos na Tabela A4 do apêndice. Percebeu-se que, no caso do Brasil, as variáveis idade do chefe, deficiência auditiva, física e mental foram as únicas que não foram estatisticamente significantes a 5%, enquanto as demais foram. Também observou-se que características como ser chefe homem, branco, possuir níveis de escolaridade como ginásio, superior e pós graduação, carteira de trabalho assinada, bem como saneamento e tratamento de lixo reduzem as chances de participar do Programa Bolsa Família.

A região do suporte comum para o Brasil foi dada entre o intervalo cujo o escore de propensão variou entre 0,09171998 e 0,99989093, no qual dentre a amostra de 193.086 famílias, 193.082 estão dentro desse intervalo, sendo um total de 128.040 tratados e 65.042 de não tratados. O teste de balanceamento¹⁶ do Brasil apontou que, após o pareamento, em quase todos os casos, houve um balanceamento moderado das variáveis, mas o pareamento por *kernel* foi o mais apropriado para fazer o *matching*. Quanto ao Pseudo-R², percebeu-se que este reduz de 0,107 para 0,001 (Tabela A5 do apêndice).

A tabela 3 mostra o Efeito Médio do Tratamento sobre os Tratados (ATT) para o Brasil e Regiões. Por meio dela, percebe-se que, em média, as famílias beneficiadas pelo programa no Brasil possuem renda do trabalho *per capita* inferior às famílias elegíveis, chegando a uma diferença de R\$19,16. Nas regiões, essa diferença é ainda maior na região Sul (-R\$20,15), enquanto a região Norte (-R\$14,73) é responsável pela menor diferença, seguida do Nordeste (-R\$16,80) e Sudeste (-R\$16,85).

Dessa forma, pode-se verificar que no Brasil e Regiões os beneficiários do Bolsa Família possuem renda inferior *per capita* quando comparados com aqueles elegíveis que não participam do programa, sendo que, para as regiões, o maior e o menor impacto ocorreu nas Região Sul e Norte, respectivamente.

Tabela 3 - Impacto do PBF na renda do trabalho *per capita*

Regiões	Tratados	Controle	ATT	teste t
Brasil	61,33	80,48	-19,16	-90,67
Nordeste	61,57	78,37	-16,80	-59,66
Norte	62,23	76,96	-14,73	-29,54
Sudeste	75,01	91,87	-16,85	-37,29
Centro-Oeste	77,84	94,91	-17,07	-18,73
Sul	72,59	92,74	-20,15	-34,46

Fonte: Elaborado pelos autores, com base nas estimações.

¹⁶ O método de Kernel também foi o mais adequado para as regiões, com exceção da região Nordeste em que o método mais apropriado foi o Vizinho mais próximo

4.3 Resultados para o Modelo de Firpo (2007)

4.3.1 Impacto do Programa Bolsa Família nas horas trabalhadas do chefe familiar

Os efeitos do Bolsa Família nas horas trabalhadas do chefe familiar para as regiões, na distribuição, são apresentados na Tabela 4. Para os *quantis*, os resultados apresentam efeitos distintos. Pode-se perceber que no primeiro *quantil*, em que a quantidade de horas trabalhadas é de 2h semanais (para o Nordeste, Norte, Sudeste e Centro-Oeste) e 3h semanais (para o Sul), não existe “efeito desincentivo” ao trabalho para os trabalhadores. A não existência desse efeito persiste nos primeiros *quantis* para a região Nordeste e Norte. Já na região Norte, é possível perceber efeitos negativos mais fortes nos *quantis* mais elevados e um pequeno impacto positivo entre os *quantis* 0,1 e 0,25.

Tabela 4 - Impacto do PBF nas horas trabalhadas dos chefes familiares – Brasil e Regiões.

Quantis	Brasil			Nordeste			Norte			Sudeste			Centro-Oeste			Sul		
	H	EQT	%	H	EQT	%	H	EQT	%	H	EQT	%	H	EQT	%	H	EQT	%
0,01	2	0 (0,135)	-	2	0 (0,176)	-	2	0 (0,284)	-	2	0 (0,412)	-	2	0 (0,840)	-	3	0 (0,479)	-
0,1	9	-1*** (0,136)	-11,1%	8	0 (0,212)	-	8	0 (0,335)	-	9	-2*** (0,502)	-22,2%	8	-2* (1,106)	-25,0%	14	-2*** (0,761)	-14,3%
0,25	20	0 (0,105)	-	20	0 (0,154)	-	20	0 (0,363)	0,0%	20	-4*** (0,873)	-20,0%	20	0 (1,307)	-	25	-8*** (0,647)	-32,0%
0,5	36	-8*** (0,260)	-22,2%	30	-6*** (0,949)	-20,0%	30	0 (0,450)	-	40	0 (0,246)	-	40	-4*** (0,971)	-10,0%	40	0 (0,242)	-
0,75	40	0 (0,043)	-	40	0 (0,064)	-	40	0 (0,172)	-	45	-1** (0,405)	-2,2%	44	-1 (0,711)	-2,3%	48	0 (0,372)	-
0,95	60	-4*** (0,278)	-6,7%	50	-4*** (1,264)	-8,0%	58	-4*** (0,912)	-6,9%	60	0 (0,730)	-	60	0 (1,038)	-	65	-5** (2,359)	-7,7%
0,99	77	-8*** (0,705)	-10,4%	72	0 (0,569)	-	80	-10*** (1,786)	-12,5%	80	-6*** (2,130)	-7,5%	80	-12* (4,642)	-15,0%	80	0 (1,838)	-

Nota: (¹) Os valores entre parênteses representam o desvio-padrão. (²) A tabela apresenta somente alguns quantis selecionados. (³) ***Estatisticamente significativo a 1%, **Estatisticamente significativo a 5%, *Estatisticamente significativo a 10%. Fonte: Elaborado pelos autores, com base nas estimações

A região Centro-Oeste tem a maior redução (12h) na jornada de trabalho, mais precisamente para os indivíduos que trabalham 80h semanais. Para as regiões Sudeste e Sul, destacam-se efeitos negativos entre os *quantis* 0,1 e 0,25. Ademais, a região Centro-Oeste é a que possui maiores efeitos desincentivo ao trabalho, enquanto a região Norte é responsável pelos menores efeitos. Para aqueles que trabalham 40h semanais tanto para o Brasil quanto para as regiões, não existem efeitos significativos, com exceção da região Centro-Oeste, em que parte dos beneficiários que se encontram nessa faixa chegam a ofertar até 4h a menos com relação aos que não são beneficiários. Estes resultados seguem os encontrados por Costa *et al* (2014), os quais consideraram para o Brasil urbano e rural.

Destacam-se alguns efeitos encontrados: 1) Efeitos neutros, não significantes, para os *quantis* em que são poucas as horas trabalhadas do chefe familiar e para os que trabalham 40h semanais nas regiões, com exceção da Centro-Oeste. 2) Efeitos negativos para os que trabalham acima de 40h semanais em todas as regiões, excetuando-se os *quantis* com efeito nulo, neste caso, esta redução nas horas trabalhadas pode ser ocasionada pelo fato desses indivíduos não quererem ultrapassar a linha de pobreza estipulada pelo Governo, e assim poder continuar recebendo os benefícios do programa, pois presume-se que quanto maior a carga horária de trabalho ofertada, maiores seriam os rendimentos. 3) Efeitos positivos entre os *quantis* 0,1 e 0,25 da região Norte, no qual as famílias beneficiadas encontram incentivos para aumentar a quantidade de horas trabalhadas.

4.3.2 Impacto do Programa Bolsa Família na renda do trabalho per capita

A Tabela 5 retrata o impacto do Programa Bolsa Família na renda do trabalho *per capita*. Por meio dela, percebe-se que, à exceção do primeiro *quantil* da região Nordeste, a renda dos beneficiários do PBF, comparada aos não beneficiários, é menor ao longo de toda a distribuição para todas as regiões do país.

Tabela 5 – Impacto do PBF na Renda do trabalho *per capita* – Regiões.

Quantis	Brasil			Nordeste			Norte			Sudeste			Centro- Oeste			Sul		
	RS	EQT	%	RS	EQT	%	RS	EQT	%	RS	EQT	%	RS	EQT	%	RS	EQT	%
0,01	6,67	-1,50*** (0,231)	-22,5%	5,55	0,00 (0,325)	0,0%	8,33	-1,09* (0,580)	-13,1%	12,00	-2,50** (1,097)	-20,8%	12,50	-4,29** (1,909)	-34,3%	12,00	-2,50*** (0,938)	-20,8%
0,1	20,00	-5,00*** (0,293)	-25,0%	17,50	-2,26*** (0,339)	-12,9%	25,00	-5,56*** (0,633)	-22,2%	35,71	-6,67*** (1,028)	-18,7%	40,00	-13,20*** (2,023)	-33,0%	33,33	-11,67*** (1,212)	-35,0%
0,25	40,00	-14,29*** (0,295)	-35,7%	33,33	-4,63*** (0,435)	-13,9%	42,86	-10,00*** (0,731)	-23,3%	60,00	-16,67*** (0,912)	-27,8%	66,67	-15,00*** (1,760)	-22,5%	59,20	-16,67*** (1,042)	-28,2%
0,5	67,00	-18,75*** (0,376)	-28,0%	60,00	-11,67*** (0,562)	-19,4%	70,38	-15,83*** (0,805)	-22,5%	87,50	-20,00*** (0,519)	-22,9%	100,00	-15,00*** (1,295)	-15,0%	85,72	-25,00*** (0,755)	-29,2%
0,75	100,00	-22,22*** (0,629)	-22,2%	90,00	-16,67*** (0,375)	-18,5%	100,00	-14,50*** (0,628)	-14,5%	107,15	-23,00*** (0,303)	-21,5%	116,67	-25,50*** (0,572)	-21,9%	106,25	-25,00*** (0,423)	-23,5%
0,95	127,50	-20,83*** (0,262)	-16,3%	125,00	-22,75*** (0,202)	-18,2%	127,50	-23,33*** (0,746)	-18,3%	127,50	-5,83*** (0,172)	-4,6%	132,00	-5,83*** (0,396)	-4,4%	130,00	-8,33*** (0,537)	-6,4%
0,99	134,50	-12,5*** (0,043)	-9,3%	133,33	-13,08*** (0,432)	-9,8%	135,71	-14,29*** (0,631)	-10,5%	138,33	-12,50*** (0,064)	-9,0%	140,00	-12,50*** (0,232)	-8,9%	140,00	-12,50*** (0,154)	-8,9%

Nota: (¹) Os valores entre parênteses representam o desvio-padrão. (²) A tabela apresenta somente alguns quantis selecionados. (³) ***Estatisticamente significativa a 1%, **Estatisticamente significativa a 5%, *Estatisticamente significativa a 10%. Fonte: Elaborado pelos autores, com base nas estimações.

Com relação às regiões Nordeste e Norte, os maiores impactos em termos monetários são observados da mediana pra cima, chegando a receber até R\$22,75 (Nordeste) e R\$23,33 (Norte) a menos que as famílias elegíveis não beneficiadas. Já nas regiões Sudeste, Centro-Oeste e Sul, a redução na renda é maior no *quantil* 0,75 (assim como para o Brasil), correspondendo a uma diminuição de R\$23,00, R\$25,50 e R\$25,00, respectivamente.

Tratando-se do impacto percentual no *quantil*, os efeitos negativos são maior encontrados nas regiões Sul e Centro-Oeste, no qual os chefes domiciliares beneficiados recebem 35,0% (*quantil* 0,1) e 34,3% (*quantil* 0,01), respectivamente, a menos que os não beneficiados. Dessa forma, embora haja uma menor renda oriunda do trabalho dos beneficiários comparado aos não beneficiários, não se pode afirmar que essa diferença seja atribuída exclusivamente ao efeito desincentivo ao trabalho, visto que foram observados distintos efeitos na distribuição.

CONSIDERAÇÕES FINAIS

Este trabalho procurou avaliar possíveis efeitos do Programa Bolsa Família nas horas trabalhadas do chefe familiar e na renda do trabalho *per capita* dos indivíduos que situam-se em áreas rurais, contribuindo com a literatura em vista dos poucos estudos existentes que buscam avaliar o impacto do programa no mercado de trabalho destas áreas. A análise foi feita utilizando dados do Censo Demográfico de 2010 e duas metodologias: *Propensity Score Matching* (PSM) e de Firpo (2007). A principal diferença entre os dois métodos reside no fato de que o primeiro analisa o efeito médio nas variáveis, enquanto o segundo avalia o efeito ao longo da distribuição.

Considerando os efeitos médios, foi possível observar, em todas as regiões e Brasil, resultados negativos sobre as horas trabalhadas e a renda do trabalho dos beneficiários do programa comparado às famílias que são elegíveis, mas não são beneficiadas. Assim, constatou-se um efeito desincentivo para o Brasil e Regiões, sendo este efeito de maior impacto na região Sul, enquanto o menor impacto foi registrado na região Norte. Já o Nordeste que é a região com a maior quantidade de beneficiários do programa no Brasil, foi a segunda região com menor impacto nas horas trabalhadas, no entanto, esta região é a que oferta menos horas trabalhadas em média no país, enquanto a região sul é a que oferta mais, logo, para os beneficiários que se encontram na região Nordeste, uma maior redução das horas trabalhadas poderia implicar em uma perda monetária mais significativa comparado aos beneficiários da região sul.

Quanto aos efeitos na distribuição, foram verificados três tipos de efeitos nas horas trabalhadas do chefe familiar: 1) Efeitos neutros nos *quantis* em que as horas trabalhadas

equivalem a no máximo 3h, para todas as regiões, bem como nos *quantis* em que as horas trabalhadas é de 40h, com exceção da região Centro-Oeste e em outros *quantis* pontuais. 2) Efeitos negativos, especialmente, onde as horas trabalhadas são superiores a 40h, talvez explicada pelo fato de que os beneficiários não queiram ultrapassar o teto de rendimento estipulado para a linha de pobreza, permanecendo com o benefício. 3) Efeitos positivos entre os *quantis* 0,1 e 0,25 da região Norte. Ressalta-se ainda que os maiores e menores efeitos desincentivo ao trabalho foram observados na região Centro-Oeste e Norte, respectivamente. Para a renda do trabalho, os efeitos na distribuição foram negativos, exceto para o primeiro *quantil* da região Nordeste, ou seja os beneficiários possuem menor renda do trabalho do que os não beneficiários.

Estes resultados não necessariamente implicam que essa menor renda do trabalho *per capita* dos beneficiários é consequência da redução de horas trabalhadas do chefe familiar, visto que os efeitos negativos nas horas trabalhadas foram em poucos *quantis*, enquanto na renda foi ao longo de todos os *quantis* analisados. Diante disso, devem existir outros fatores que tenham impacto no mercado de trabalho que não foram identificados nos resultados aqui obtidos, como, por exemplo, os postos de trabalho ocupados pelos indivíduos e o diferente dinamismo econômico de cada região rural do país.

REFERÊNCIAS

AVERBUG, André. **A Economia Brasileira nos anos 90: Abertura e Integração Comercial na Década de 90**. Rio de Janeiro: BNDES, 1999.

BRASIL, Ministério do Desenvolvimento Social e Combate à Fome, 2015. Disponível em: <<http://www.mds.gov.br/bolsafamilia>>. Acesso em: 06 abr. 2015.

_____. **Sistema de Benefícios ao Cidadão – SIBEC (2015)**, Disponível em: <<http://www.mds.gov.br/bolsafamilia>>. Acesso em: 06 abr. 2015.

BRITO, A. S. **Programa Bolsa Família e Mercado de Trabalho: Uma Análise das Limitações e Possibilidades da PNAD e do Cadastro Único**. 2011. Dissertação. (Mestrado). Universidade Federal Fluminense, Rio de Janeiro, 2011.

BRITO, D. J. M. de; ROCHA, R. M. Concentração de renda e escolaridade: qual é a relação com a pobreza no Nordeste do país?. **Bahia Análise & Dados**, Bahia, v. 21, p. 583-594, 2011.

CAMPÊLO, A. K.; SILVA, E. N. Filhos e renda familiar: uma aplicação do efeito quantílico de tratamento. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 35, n. 3, p. 355-377, dez. 2005

CARDOSO, E.; SOUZA, A. P. **The impact of cash transfers on child labor and school attendance in Brazil**. Nashville: Department of Economics, Vanderbilt University, 2004. (Working Paper 04-W07). Disponível em: <<http://discoverarchive.vanderbilt.edu/bitstream/handle/1803/20/vu04-w07.pdf?sequence=1>>. Acesso em: 06 abr. 2015

CARVALHO, R.; MARINHO, E. L. L. Transformações Estruturais, Variações na Eficiência Técnica e Produtividade Total dos Fatores no Setor Agrícola dos Países Sul-Americanos – 1970 a 2000. In: ECONTRO NACIONAL DE ECONOMIA, 31., 2003. **Anais...** Porto Seguro, ANPEC, 2003.

CAVALCANTI, D. M. **Avaliação dos Impactos do Programa Bolsa Família na Renda, Educação e no Mercado de Trabalho das Famílias Pobres do Brasil**. 2013. Dissertação. (Mestrado). Universidade Federal do Rio Grande do Norte, Programa de Pós-Graduação em Economia, Centro de Ciências Sociais Aplicadas. 2013.

CECCHINI, S. Transferências condicionadas na América Latina e Caribe: da inovação à consolidação. In: Campello, T.; NERI, M. C. (Eds.). **Programa Bolsa Família: uma década de inclusão e cidadania**. Brasília: Ipea, 2013.

CEDEPLAR – Centro de Desenvolvimento e Planejamento Regional. **Projeto de avaliação do impacto do programa Bolsa Família** – relatório analítico final. 2006. Mimeografado.

COSTA, E. W.; CAVALCANTI, D. M.; SILVA, J. L. M. da; SAMPAIO, R. M. B. Impactos Regionais do Programa Bolsa Família: Uma Análise sobre o Mercado de Trabalho das Famílias Pobres. *In*: ANPEC: XLII ENCONTRO NACIONAL DE ECONOMIA, 2014. Disponível em: < http://www.anpec.org.br/encontro/2014/submissao/files_I/i13-dc0aad69d0a6580ef400e1ea2c1accc7.pdf> Acesso em: 20 abr. 2015.

FERREIRA, M. de O.; RAMOS, L. M.; ROSA, A. L. T. da; LIMA, P. V. P. S.; LEITE, L. A. de S. Especialização produtiva e mudança estrutural da agropecuária cearense. **Teoria e Evidência Econômica**, Passo Fundo, v. 14, n. 26, p. 91-111, maio 2006.

FIRPO, Sérgio. Efficient Semiparametric Estimation of *Quantile* Treatment Effects. **Econométrica**, vol. 75, n.1, 2007.

FOGUEL, M.; BARROS, R. P. de. The effects of conditional cash transfer programmes on adult labour supply: an empirical analysis using a time-series-cross-section sample of Brazilian municipalities. **Estudos Econômicos**, v. 40, n. 2, p. 259-293, 2010.

IBGE – Instituto Brasileiro de Geografia e Estatística. Banco de Dados Agregados. **Sistema IBGE de Recuperação Automática - SIDRA**. Disponível em: <<http://www.ibge.gov.br>> Acesso em: 20 abr. 2015.

_____. Censo Demográfico de 2010. Rio de Janeiro. Disponível em: <<http://www.ibge.gov.br>>. Vários acessos.

_____. Pesquisa Nacional por Amostra de Domicílios de 2013. Rio de Janeiro. Disponível em: <<http://www.ibge.gov.br>>. Vários acessos.

KHANDKER, S.; KOOLWAL, G; SAMAD, H. **Handbook on Impact Evaluation: Quantitative Methods and Practices**. World Bank, October 2009.

LEVY, S. **Buenas intenciones, malos resultados**: política social, informalidad y crecimiento económico en México. Mexico: Océano, 2010.

MATTOS, E; MAIA, S. e MARQUES, F. Evidências da relação entre oferta de trabalho e programas de transferência de renda no Brasil: bolsa escola versus renda mínima. **Pesquisa e Planejamento Econômico**. Brasília: v. 40, n. 2, ago. 2010.

NASCIMENTO, C. A.; OLIVEIRA, R. B; SOUTO, I. J.G e MENDES, S. R. A Qualidade do emprego rural fluminense nos anos recentes (2002 e 2006). *In*: CONGRESSO DA SOCIEDADE BRASILEIRA DE ECONOMIA E SOCIOLOGIA RURAL, 46, 2008, Rio Branco. **Anais...** Rio Branco, 2008. Disponível em: <<http://www.sober.org.br>> Acesso em: 05 mai. 2015.

NUNES, J. A., MARIANO, J. L. Efeitos dos Programas de Transferência de Renda Sobre a Oferta de Trabalho Não Agrícola na Área Rural da Região Nordeste. **Revista de Economia e Sociologia Rural**, 2015, v. 53, n. 01, p.71-90.

SANTOS, G. C.; FONTES, R. M. O.; BASTOS, P. M. A.; LIMA, J. E. de.. Mercado de trabalho e rendimento no meio rural brasileiro. **Economia Aplicada**, Ribeirão Preto, v. 14, n. 3, p. 355-379, set. 2010.

SOARES, S; SÁTYRO, N. **O Programa Bolsa Família: Desenho Institucional, Impactos e Possibilidades Futuras**; Brasília: Ipea, 2009 (Texto para discussão, n. 1424). Disponível em: <<http://www.ipea.gov.br>> Acesso em: 20 abr. 2015.

TAVARES, P. A. Efeito do Programa Bolsa Família sobre a oferta de trabalho das mães. *In*: XIII SEMINÁRIO SOBRE ECONOMIA MINEIRA – ECONOMIA, HISTÓRIA, DEMOGRAFIA E POLÍTICAS PÚBLICAS. Diamantina, MG, 2008.

TEIXEIRA, C. G., **Análise do Impacto do Programa Bolsa Família na oferta de trabalho dos homens e das mulheres**. PNUD/IPC, 2008. Disponível em:<<http://www.ipc-undp.org/publications/mds/27P.pdf>> Acesso em: 20 abr. 2015.

TEIXEIRA, J. C., Modernização da Agricultura no Brasil: Impactos Econômicos, Sociais e Ambientais. **Revista Eletrônica da Associação dos Geógrafos Brasileiros – Seção Três Lagoas**, v.2, n.2, 2005.

APÊNDICE

Tabela A1 – Estimação do Modelo Logit para as horas trabalhadas.

Variáveis	Brasil			Nordeste			Norte			Sudeste			Centro-Oeste			Sul		
	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z
idade_chefe	-0,004	0,001	0,000	0,011	0,001	0,000	0,001	0,002	0,582	-0,011	0,002	0,000	-0,020	0,005	0,000	-0,006	0,003	0,031
chef_homem	-0,136	0,018	0,000	-0,144	0,025	0,000	-0,071	0,046	0,126	-0,021	0,053	0,699	-0,075	0,110	0,496	-0,078	0,059	0,182
chef_branco	-0,412	0,018	0,000	-0,100	0,025	0,000	-0,130	0,054	0,016	-0,458	0,047	0,000	-0,192	0,108	0,076	-0,657	0,056	0,000
ensino_2o_grau	-0,545	0,032	0,000	-0,127	0,032	0,000	0,301	0,059	0,000	-0,223	0,068	0,001	0,166	0,137	0,224	-0,375	0,079	0,000
ensino_ginásio	-0,252	0,023	0,000	0,129	0,023	0,000	0,472	0,043	0,000	0,081	0,054	0,132	0,398	0,114	0,000	-0,125	0,070	0,073
ensino_primário	0,033	0,017	0,053	-0,405	0,043	0,000	0,155	0,086	0,070	-0,641	0,098	0,000	0,275	0,177	0,120	-0,903	0,112	0,000
ensino_superior	-1,582	0,171	0,000	-1,392	0,216	0,000	-1,437	0,532	0,007	-1,313	0,537	0,014	-1,921	1,078	0,075	-2,021	0,630	0,001
mestrado_doutorado	-2,264	0,510	0,000	-1,612	0,598	0,007	-	-	-	-	-	-	-	-	-	-	-	-
cart_trab	-0,326	0,040	0,000	-0,497	0,059	0,000	-0,163	0,140	0,246	-0,250	0,080	0,002	-0,407	0,181	0,025	-0,297	0,110	0,007
nfilho_5	0,169	0,010	0,000	0,219	0,015	0,000	0,053	0,020	0,009	0,292	0,029	0,000	0,234	0,058	0,000	0,500	0,035	0,000
nfilho_10	0,563	0,010	0,000	0,693	0,015	0,000	0,444	0,020	0,000	0,607	0,026	0,000	0,599	0,054	0,000	0,633	0,030	0,000
nfilho_15	0,542	0,009	0,000	0,586	0,014	0,000	0,468	0,021	0,000	0,599	0,025	0,000	0,628	0,053	0,000	0,678	0,029	0,000
nfilho_17	0,488	0,018	0,000	0,559	0,027	0,000	0,430	0,042	0,000	0,554	0,049	0,000	0,529	0,110	0,000	0,430	0,052	0,000
nfilhos_18	0,192	0,010	0,000	0,177	0,014	0,000	0,136	0,020	0,000	0,117	0,029	0,000	0,200	0,068	0,003	0,093	0,039	0,017
filho_branco	0,206	0,017	0,000	0,213	0,024	0,000	0,231	0,047	0,000	0,233	0,046	0,000	0,134	0,104	0,197	0,162	0,057	0,005
auditiva	0,130	0,071	0,067	0,117	0,101	0,248	-0,245	0,187	0,190	-0,084	0,197	0,671	0,579	0,464	0,212	0,338	0,189	0,073
fisica	0,021	0,054	0,690	-0,119	0,075	0,111	0,259	0,132	0,050	0,000	0,165	1,000	0,075	0,351	0,831	0,145	0,146	0,319
visual	0,185	0,034	0,000	-0,192	0,110	0,082	-0,127	0,199	0,524	0,085	0,243	0,727	0,266	0,505	0,598	0,411	0,247	0,096
mental	-0,040	0,082	0,627	0,120	0,048	0,011	0,231	0,081	0,004	0,178	0,102	0,079	-0,164	0,230	0,475	0,259	0,095	0,007
agricultor	-0,148	0,017	0,000	-0,043	0,024	0,072	-0,284	0,045	0,000	0,103	0,049	0,036	-0,239	0,105	0,022	-0,217	0,057	0,000
condi_dom	0,094	0,018	0,000	0,136	0,026	0,000	0,143	0,055	0,009	0,103	0,042	0,015	0,068	0,094	0,471	-0,091	0,051	0,075
material_dom	0,529	0,016	0,000	0,286	0,025	0,000	0,095	0,047	0,043	-0,104	0,076	0,172	0,089	0,102	0,384	-0,363	0,043	0,000
tipo_dom	0,745	0,075	0,000	0,505	0,144	0,000	1,147	0,145	0,000	0,254	0,250	0,308	0,527	0,203	0,009	0,212	0,233	0,363
saneamento	-0,409	0,023	0,000	-0,180	0,037	0,000	-0,373	0,085	0,000	-0,266	0,055	0,000	-0,263	0,145	0,071	-0,263	0,051	0,000
agua_canalizada	0,358	0,017	0,000	0,081	0,023	0,000	0,373	0,055	0,000	0,445	0,057	0,000	0,424	0,121	0,000	0,448	0,046	0,000
tratam_lixo	-0,379	0,021	0,000	-0,134	0,031	0,000	-0,218	0,075	0,004	-0,493	0,052	0,000	-0,315	0,153	0,039	-0,386	0,048	0,000
eletricidade	0,403	0,019	0,000	0,316	0,027	0,000	0,420	0,040	0,000	0,077	0,060	0,204	0,061	0,124	0,620	-0,181	0,077	0,019
_cons	-1,449	0,083	0,000	-1,475	0,151	0,000	-2,610	0,174	0,000	-0,522	0,265	0,049	-1,409	0,304	0,000	-0,165	0,254	0,517
Log likelihood	-60043,60			-30854,34			-9108,22			-8133,32			-1644,07			-7313,50		
Pseudo R2	0,12920			0,12620			0,1222			0,1314			0,1534			0,1588		
Nº de Observ.	103.648			59.284			14.990			13.695			2.906			12.756		
Suporte Comum	[0.07688124, 0.99987008]			[0.11227859, 0.99910881]			[0.058789, 0.9711901]			[0.10045646, 0.99980859]			[0.04481351, 0.99596426]			[0.07296871, 0.99399378]		

Fonte: Elaboração própria

Tabela A3 – Testes de comparação das médias dos grupos dos tratados e controle, antes e após os pareamentos – horas trabalhadas – Brasil.

Variáveis	Antes do Pareamento			Após pareamento								
				Vizinho mais próximo			Alcance			Kernel		
	Tratados	Controle	p valor	Tratados	Controle	p valor	Tratados	Controle	p valor	Tratados	Controle	p valor
idade_chefe	39,46	38,68	0,000	38,94	38,68	0,002	39,46	39,49	0,606	39,46	39,488	0,631
chef_homem	0,704	0,746	0,000	0,733	0,746	0,000	0,704	0,705	0,664	0,704	0,705	0,657
chef_branco	0,273	0,371	0,000	0,338	0,371	0,000	0,273	0,273	0,872	0,273	0,273	0,840
ensino_ginasio	0,143	0,196	0,000	0,179	0,196	0,000	0,143	0,139	0,060	0,143	0,139	0,061
ensino_primario	0,535	0,480	0,000	0,508	0,480	0,000	0,535	0,530	0,076	0,535	0,530	0,800
ensino_2o_grau	0,043	0,084	0,000	0,062	0,084	0,000	0,043	0,043	0,914	0,043	0,043	0,901
ensino_superior	0,001	0,004	0,000	0,001	0,004	0,000	0,001	0,001	0,559	0,001	0,001	0,572
mestrado_doutorado	7,8E-05	0,001	0,000	0,000	0,001	0,000	7,8E-05	0,000	0,746	7,8E-05	0,000	0,733
cart_trab	0,035	0,032	0,006	0,034	0,032	0,000	0,035	0,037	0,043	0,035	0,037	0,046
nfilho_5	0,609	0,588	0,000	0,613	0,588	0,000	0,609	0,655	0,000	0,609	0,655	0,000
nfilho_10	0,828	0,427	0,000	0,513	0,427	0,000	0,828	0,853	0,000	0,828	0,854	0,000
nfilho_15	0,860	0,396	0,000	0,455	0,396	0,000	0,860	0,855	0,374	0,860	0,856	0,388
nfilho_17	0,252	0,125	0,000	0,149	0,125	0,000	0,252	0,248	0,184	0,252	0,248	0,185
nfilhos_18	0,450	0,326	0,000	0,380	0,326	0,000	0,450	0,470	0,000	0,450	0,470	0,000
filho_branco	0,402	0,361	0,000	0,388	0,361	0,041	0,402	0,407	0,082	0,402	0,407	0,080
auditiva	0,011	0,009	0,045	0,010	0,009	0,537	0,011	0,011	0,682	0,011	0,011	0,724
fisica	0,018	0,018	0,451	0,018	0,018	0,085	0,018	0,019	0,402	0,018	0,019	0,387
mental	0,007	0,008	0,416	0,007	0,008	0,936	0,007	0,007	0,761	0,007	0,007	0,768
visual	0,053	0,045	0,000	0,048	0,045	0,798	0,053	0,053	0,638	0,053	0,053	0,617
agro_chef	0,666	0,721	0,000	0,698	0,721	0,004	0,666	0,658	0,001	0,666	0,658	0,001
condi_dom	0,816	0,790	0,000	0,799	0,790	0,000	0,816	0,808	0,000	0,816	0,808	0,000
material_dom	0,715	0,612	0,000	0,666	0,612	0,000	0,715	0,694	0,000	0,715	0,694	0,000
tipo_dom	0,994	0,984	0,000	0,992	0,984	0,000	0,994	0,992	0,000	0,994	0,992	0,000
saneamento	0,086	0,120	0,000	0,105	0,120	0,000	0,086	0,088	0,146	0,086	0,088	0,155
agua_canalizada	0,281	0,218	0,000	0,240	0,218	0,000	0,281	0,279	0,460	0,281	0,279	0,486
tratam_lixo	0,132	0,160	0,000	0,150	0,160	0,000	0,132	0,136	0,023	0,132	0,136	0,024
eletricidade	0,822	0,753	0,000	0,790	0,753	0,000	0,822	0,801	0,000	0,822	0,801	0,000
Pseudo-R2	0,126			0,018			0,002			0,002		
LR chi2	17415,31			1967,76			296,63			299,2		

Fonte: Elaborado pelos autores, com base nas estimações

Tabela A4 – Estimação do Modelo Logit para a renda do trabalho *per capita*.

Variáveis	Brasil			Nordeste			Norte			Sudeste			Centro-Oeste			Sul		
	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z	Coef.	Desvio-Padrão	P>z
idade_chefe	0,000	0,001	0,698	0,008	0,001	0,000	0,007	0,002	0,000	-0,003	0,002	0,052	-0,005	0,003	0,128	-0,004	0,002	0,031
chef_homem	-0,247	0,014	0,000	-0,247	0,019	0,000	-0,226	0,038	0,000	-0,150	0,039	0,000	-0,196	0,079	0,013	-0,178	0,048	0,000
chef_branco	-0,276	0,013	0,000	-0,077	0,019	0,000	-0,126	0,040	0,002	-0,329	0,032	0,000	-0,179	0,072	0,013	-0,419	0,045	0,000
ensino_2o_grau	-0,444	0,023	0,000	-0,118	0,023	0,000	0,225	0,044	0,000	-0,096	0,045	0,033	0,111	0,088	0,207	-0,300	0,062	0,000
ensino_ginásio	-0,200	0,017	0,000	0,129	0,017	0,000	0,317	0,032	0,000	0,131	0,035	0,000	0,256	0,071	0,000	-0,062	0,053	0,243
ensino_primário	0,036	0,013	0,004	-0,360	0,031	0,000	0,063	0,062	0,312	-0,458	0,067	0,000	0,058	0,123	0,639	-0,712	0,087	0,000
ensino_superior	-1,334	0,120	0,000	-1,252	0,159	0,000	-1,141	0,338	0,001	-1,044	0,366	0,004	-1,087	0,453	0,016	-1,288	0,429	0,003
mestrado_doutorado	-1,450	0,335	0,000	-0,726	0,477	0,128	-2,128	0,773	0,006	-1,514	1,134	0,182	-	-	-	-	-	-
cart_trab	-0,686	0,018	0,000	-0,799	0,028	0,000	-0,465	0,063	0,000	-0,529	0,036	0,000	-0,632	0,081	0,000	-0,619	0,058	0,000
nfilho_5	0,185	0,007	0,000	0,260	0,011	0,000	0,070	0,016	0,000	0,250	0,018	0,000	0,173	0,037	0,000	0,352	0,026	0,000
nfilho_10	0,554	0,007	0,000	0,688	0,011	0,000	0,453	0,016	0,000	0,586	0,017	0,000	0,541	0,035	0,000	0,548	0,024	0,000
nfilho_15	0,505	0,007	0,000	0,595	0,011	0,000	0,432	0,016	0,000	0,521	0,017	0,000	0,487	0,035	0,000	0,548	0,023	0,000
nfilho_17	0,470	0,014	0,000	0,546	0,020	0,000	0,419	0,033	0,000	0,501	0,033	0,000	0,479	0,070	0,000	0,369	0,043	0,000
nfilhos_18	0,149	0,007	0,000	0,159	0,010	0,000	0,073	0,015	0,000	0,103	0,020	0,000	0,117	0,045	0,008	0,078	0,032	0,015
filho_branco	0,208	0,013	0,000	0,247	0,018	0,000	0,197	0,035	0,000	0,139	0,032	0,000	0,121	0,068	0,077	0,169	0,046	0,000
auditiva	0,024	0,053	0,648	-0,045	0,072	0,529	-0,067	0,143	0,640	-0,031	0,137	0,820	0,332	0,290	0,253	0,216	0,160	0,178
fisica	0,005	0,042	0,907	-0,072	0,057	0,208	0,261	0,106	0,013	-0,132	0,114	0,247	0,075	0,272	0,782	0,173	0,123	0,161
visual	0,146	0,026	0,000	-0,179	0,086	0,036	0,048	0,166	0,772	0,018	0,169	0,916	0,872	0,428	0,042	0,162	0,194	0,404
mental	-0,041	0,064	0,517	0,098	0,036	0,006	0,156	0,063	0,014	0,213	0,074	0,004	0,058	0,154	0,707	0,101	0,081	0,212
agricultor	0,133	0,012	0,000	0,190	0,016	0,000	0,032	0,032	0,311	0,240	0,031	0,000	-0,042	0,064	0,508	0,114	0,042	0,006
condi_dom	0,302	0,012	0,000	0,311	0,017	0,000	0,270	0,037	0,000	0,256	0,028	0,000	0,199	0,061	0,001	0,133	0,039	0,001
material_dom	0,430	0,012	0,000	0,245	0,019	0,000	0,142	0,034	0,000	-0,073	0,052	0,157	0,082	0,065	0,207	-0,215	0,036	0,000
tipo_dom	0,222	0,057	0,000	0,133	0,107	0,211	0,393	0,121	0,001	0,196	0,163	0,228	-0,082	0,132	0,534	0,533	0,165	0,001
saneamento	-0,347	0,017	0,000	-0,208	0,026	0,000	-0,337	0,059	0,000	-0,289	0,035	0,000	-0,209	0,092	0,023	-0,180	0,043	0,000
agua_canalizada	0,298	0,013	0,000	0,077	0,017	0,000	0,266	0,040	0,000	0,340	0,037	0,000	0,296	0,081	0,000	0,358	0,039	0,000
tratam_lixo	-0,379	0,015	0,000	-0,255	0,021	0,000	-0,237	0,049	0,000	-0,440	0,032	0,000	-0,202	0,094	0,031	-0,280	0,040	0,000
eletricidade	0,367	0,014	0,000	0,334	0,019	0,000	0,317	0,030	0,000	0,110	0,041	0,006	0,168	0,079	0,034	0,078	0,057	0,169
_cons	-0,999	0,062	0,000	-1,002	0,111	0,000	-1,769	0,141	0,000	-0,756	0,173	0,000	-1,004	0,198	0,000	-0,776	0,185	0,000
Log likelihood	-109832,59			-58389,23			15726,34			-17931,14			-3788,74			-10364,87		
Pseudo R2	0,10970			0,1242			0,0976			0,1064			0,0929			0,0984		
Nº de Observ.	193.086			115.475			25.312			29.559			6.049			16.688		
SupORTE Comum	[0.09171998, 0.99989093]			[0.11194566, 0.99979729]			[0.03859959, 0.99165672]			[0.098113, 0.99979052]			[0.07783643, 0.98841281]			[0.10334952, 0.98705872]		

Fonte: Elaborado pelos autores, com base nas estimações

Tabela A5 – Testes de comparação das médias dos grupos dos tratados e controle, antes e após os pareamentos – renda do trabalho - Brasil.

Variáveis	Antes do Pareamento			Após pareamento								
				Vizinho mais próximo			Alcance			Kernel		
	Tratados	Controle	p valor	Tratados	Controle	p valor	Tratados	Controle	p valor	Tratados	Controle	p valor
idade_chefe	39,65	38,25	0,000	38,71	38,25	0,000	39,65	39,96	0,000	39,65	39,930	0,000
chef_homem	0,773	0,813	0,000	0,808	0,813	0,027	0,773	0,769	0,017	0,773	0,769	0,015
chef_branco	0,269	0,321	0,000	0,324	0,321	0,280	0,269	0,277	0,000	0,269	0,277	0,000
ensino_ginasio	0,141	0,190	0,000	0,187	0,190	0,180	0,141	0,140	0,224	0,141	0,140	0,437
ensino_primario	0,524	0,470	0,000	0,483	0,470	0,000	0,524	0,522	0,246	0,524	0,522	0,362
ensino_2o_grau	0,044	0,082	0,000	0,072	0,082	0,000	0,044	0,045	0,459	0,044	0,045	0,322
ensino_superior	0,001	0,003	0,000	0,002	0,003	0,000	0,001	0,001	0,833	0,001	0,001	0,869
mestrado_doutorado	1,2E-04	0,000	0,000	0,000	0,000	0,022	1,2E-04	0,000	0,892	1,2E-04	0,000	0,928
cart_trab	0,076	0,117	0,000	0,107	0,117	0,000	0,076	0,078	0,029	0,076	0,078	0,041
nfilho_5	0,632	0,642	0,017	0,664	0,642	0,000	0,632	0,655	0,000	0,632	0,656	0,000
nfilho_10	0,800	0,441	0,000	0,416	0,441	0,000	0,800	0,817	0,000	0,800	0,817	0,000
nfilho_15	0,818	0,404	0,000	0,345	0,404	0,000	0,818	0,819	0,644	0,818	0,819	0,626
nfilho_17	0,239	0,124	0,000	0,112	0,124	0,000	0,239	0,238	0,685	0,239	0,238	0,653
nfilhos_18	0,462	0,353	0,000	0,391	0,353	0,000	0,462	0,488	0,000	0,462	0,488	0,000
filho_branco	0,394	0,341	0,000	0,353	0,341	0,000	0,394	0,401	0,000	0,394	0,401	0,000
auditiva	0,010	0,010	0,545	0,010	0,010	0,933	0,010	0,010	0,790	0,010	0,010	0,706
fisica	0,017	0,016	0,726	0,017	0,016	0,279	0,017	0,017	0,614	0,017	0,017	0,701
mental	0,006	0,007	0,246	0,007	0,007	0,476	0,006	0,006	0,309	0,006	0,006	0,261
visual	0,049	0,042	0,000	0,043	0,042	0,320	0,049	0,049	0,529	0,049	0,049	0,632
agro_chef	0,661	0,647	0,000	0,649	0,647	0,531	0,661	0,651	0,000	0,661	0,651	0,000
condi_dom	0,786	0,709	0,000	0,728	0,709	0,000	0,786	0,783	0,083	0,786	0,783	0,079
material_dom	0,730	0,661	0,000	0,678	0,661	0,000	0,730	0,715	0,000	0,730	0,715	0,000
tipo_dom	0,994	0,988	0,000	0,991	0,988	0,000	0,994	0,992	0,000	0,994	0,992	0,000
saneamento	0,092	0,124	0,000	0,121	0,124	0,094	0,092	0,097	0,000	0,092	0,097	0,000
agua_canalizada	0,287	0,241	0,000	0,244	0,241	0,159	0,287	0,284	0,107	0,287	0,284	0,091
tratam_lixo	0,156	0,198	0,000	0,195	0,198	0,222	0,156	0,164	0,000	0,156	0,164	0,000
eletricidade	0,826	0,763	0,000	0,784	0,763	0,000	0,826	0,816	0,000	0,826	0,816	0,000
Pseudo-R2	0,107			0,004			0,001			0,001		
LR chi2	26406,32			726,93			424,67			410,84		

Fonte: Elaborado pelos autores, com base nas estimações.

Heat Shocks, Child Endowments, and Parental investments

Bladimir Carrillo
(Corresponding author)
Federal University of Vicosa
AV PH Rolfs
Viçosa, Brazil 36570
bladimir.bermudez@ufv.br

João Lima
Federal University of Vicosa
AV PH Rolfs
Viçosa, Brazil 36570
jelima@ufv.br

Juan C. Trujillo
University of York
Heslington
York, UK YO10
jct523@york.ac.uk

RESUMO

Este artigo investiga como a exposição pré-natal a temperaturas extremamente altas afeta o investimento em capital humano na Colômbia. Usando uma estratégia de efeito-fixos, encontra-se que crianças expostas a estresse de calor durante o segundo trimestre da gravidez são mais propensas a receber as vacinas necessárias e são amamentados por mais tempo. Uma variedade de evidência é apresentada em favor da interpretação de que esses resultados refletem respostas compensatórias das famílias a mudanças nas dotações iniciais das crianças. O artigo também encontra evidência sugestiva de que os efeitos de estresse de calor pré-natal podem persistir até a vida adulta, o que indica que os aumentos nos investimentos durante a infância não remediam completamente os efeitos de linha base de choques pré-natais de calor.

Palavras-Chave: Eventos extremos de temperatura; dotações iniciais das crianças; Investimentos em capital humano

ABSTRACT

This paper studies how *in utero* exposure to extreme hot temperatures affects parental investments in Colombia. Using a sibling-fixed effects strategy, we show that children who were exposed *in utero* to heat stress during second trimester are more likely to receive necessary vaccines and are breastfed for longer. A variety of evidence is presented in favor of the interpretation that this household behavior reflects a compensatory response to shifts in children's endowments. We also provide suggestive evidence that the effects of prenatal heat stress can last into adulthood, indicating that increased investments during infancy are not sufficient to remedy the baseline effects of *in utero* shocks.

Keywords: Extreme weather events; birth endowments; parental investments.

JEL Codes: D1, I1, J1

ÁREA DE SUBMISSÃO: Economia Social e do Trabalho

1. Introduction

Poor environmental conditions *in utero* can have significant effects on later-life outcomes. Compared to their counterparts, children endowed with poor health are more likely to have lower cognitive abilities, lower educational attainment, and worse health outcomes as adults (Currie and Hyson 1999; Behrman and Rosenzweig 2004; Almond 2006). One intriguing question is how parental behavior responds to shifts in initial endowments. Such parental responses indicate by how much household behavior exacerbate or mitigate the effects of *in utero* shocks. This question has been subject of theoretical debate for a long time. According to Becker and Tomes (1976), if low initial endowment in a child implies lower returns on investments, then parents are likely to adopt reinforcing strategies. Conversely, Behrman et al. (1982) conjecture that when parents care about sibling inequality in welfare, then they would make more human capital investments in the less-endowed child. Ultimately, how parental behavior responds to shifts in initial endowments is an empirical question.

In this paper, we estimate the impact of prenatal exposure to heat waves on health investments using Colombian data. Exposure to extreme high temperatures is believed to be hazardous for health. A pregnant woman is more susceptible to high temperatures due to the additional physical strain and the reduced capacity to lose heat by sweating (Strand, Barnett, and Tong 2011; Wells and Cole 2002). As a result, it has been hypothesized that exposure to heat stress during pregnancy could harm the development of the fetus. Randomized experiments with animals suggest that exogenous exposure to high temperatures *in utero* negatively affects health and the motor development of offspring (Strong et al. 2015; Shiota and Kayamura 1989). A set of recent research in humans also has found suggestive evidence that maternal heat stress during pregnancy has adverse consequences on infant health (Deschenes, Greenstone, and Guryan 2009; Strand, Barnett, and Tong 2011). Altogether, these facts suggest that greater exposure *in utero* to heat waves has the potential to explain low levels of human capital.

The question is especially relevant in view of projections indicating that extreme temperature episodes will increase in the next decades (IPCC 2007). A growing body of recent work has tried to quantify the effects of such climatic shocks on several dimensions, including health and income (Dell, Jones, and Olken 2012; Deschênes and Moretti 2009). However, most of studies focus on the short-term impacts of extreme temperatures. Remarkably, the pathways on how weather events could have long-term impacts are not well-studied. Understanding the parental responses to shifts in endowments induced by prenatal extreme hot temperatures would provide important insights.

Our identification strategy exploits plausibly exogenous variation in temperature over time within municipalities. We construct a municipality-by-month weather dataset, which then is combined with microdata by using date and place of birth to identify the prevailing temperature conditions during pregnancy. The empirical approach then compares parental investments on children that were prenatally exposed to extreme hot temperatures relative to those that experienced less extreme temperature conditions *in utero*. Since the occurrence of a temperature shock at a given moment in time and place is unpredictable, prenatal exposure to heat waves can be considered as good as randomly assigned. In addition, we can control for sibling-fixed effects to address the issue that different types of families may change their fertility decisions based on temperature around the time of conception (Barreca, Deschenes, and Guldi 2015; Wilde, Apouey, and Jung, *Forthcoming*). This research design is particularly suitable for the Colombian context. As hydro-meteorological patterns are affected by a recurrent climatic event, temperature records in Colombia vary widely year to year across municipalities. Furthermore, since agricultural production account for a low share of Colombia GDP, the potential for general equilibrium effects is diminished.

The inputs that we have examined are breastfeeding and vaccination. Both investments are believed to be particularly important for child health in developing countries. Vaccinations such as polio and measles have been shown to be effective in preventing ill health and mortality. Likewise, breastfeeding plays a central role in nutrition, especially in environments characterized by unsafe drinking water and limited supply of food. A large body of work has also documented that breastfeeding is predictive of later cognitive outcomes.¹

Using the sibling-fixed effects strategy, we document that children who were exposed to heat waves while *in utero* are more likely to receive necessary vaccines and are breastfed for longer. These results are not driven by time-series correlation in temperature, migration, or changes in local economic activity. Furthermore, we find that the quantity and spacing of births are not significantly affected. Therefore, we believe that it may be reasonable to attribute the effects to variations in child endowments. This interpretation is made somewhat more plausible by the evidence that prenatal heat stress has adverse consequences on offspring endowments and by the documented relationship, both theoretically and empirically, between child endowments and parental investments. As such, our findings point out that birth endowments are an unexamined mechanism by which extreme hot temperatures could have long-term impacts on human capital accumulation of cohorts prenatally exposed during these periods.

One implication of these findings is that the long-term impacts of prenatal heat waves on human capital accumulation are not trivial. If early health investments play an important role in the later-life capabilities, the net long-term impact of prenatal hot temperature could become small or even translated into improved human capital accumulation. To shed lights on this issue, we provide reduced-form estimates of the impacts of heat waves on total years of schooling for adults aged 20-36, using census data. We find a small, but statistically significant effect of prenatal exposure to heat waves on educational attainment. Cohorts exposed prenatally to one heat wave received an average of 0.03 fewer years (0.34 per cent) of education. This suggests that the initial shock is not completely offset by the increased health investments during infancy.

Our study contributes to a growing body of knowledge on the links between *in-utero* exposure to environmental shocks and human capital investments (Almond, Edlund, and Palme 2009; Kelly 2011; Venkataramani 2012; Parman 2013). While studies in this area typically focus on uncommon and severe historical events, we focus on an environmental shock that is less drastic but occurs with higher frequency. We see our results as a first attempt to show the systematic importance of heat stress *in utero* on parental investments. Furthermore, most of existing historical studies use limited measures of investments and have used a variety of indirect strategies to infer parental responses. For example, Almond, Edlund, and Palme (2009) argue that parents adopt reinforcing strategies because the effect of fetal exposure to the radioactive fallout on cognitive skills was greater in children from poor families. This evidence is compelling, but not conclusive. In contrast to these studies, we use more direct measures of parental inputs. Our study also builds on the recent work by Adhvaryu and Nyshadham (2014), who find that children with higher exposure to an iodine supplementation program during pregnancy received more health investments in Tanzania.

This paper is also related to a number of previous studies that link parental investments to proxy variables for endowment, such as birth weight. This literature is not conclusive.² While some studies find evidence for reinforcement (Datar, Kilburn, and Loughran 2010; Aizer and Cunha 2012), others find that parents respond with compensating behavior (Del Bono, Ermisch, and Francesconi 2012). In part, clear stylized facts are not developed due to the endogeneity issues. Prenatal and postnatal unobserved investments could create a correlation between birth endowments and parental investment, even in the absence of a

¹ See, for example, Del Bono and Rabe (2012), Anderson, Johnstone, Remley (1999), and Victora et al (2015).

² See Almond and Mazumder (2013) for an inventory.

behavior response.³ Our strategy adds to this discussion by using exposure to heat waves as a source of variation. Given the idiosyncrasy of the occurrence of a temperature shock, we argue that prenatal exposure to heat waves is more likely to be free of the endogeneity issues that plague these previous works.

The rest of the paper is structured as follows. In section 2, we provide background information on the relationship between heat stress and offspring outcomes. In sections 3 and 4, we describe our data and empirical strategy, respectively. In section 5, we present our empirical findings, including robustness checks and section 6 includes an exploration of potential long-run impacts on human capital accumulation. Section 7 concludes.

2. Background on Heat-Stress and Endowments

Exposure to high temperatures is one of the most encountered physical stressful events. Medical literature indicates that prenatal heat stress increases mother's levels of cortisol, a hormone that plays a critical role in fetal health and the regulation of the psychomotor development (Davis and Sandman 2010; Wadhwa et al. 1993).⁴ An early study by Vaha-Eskeli et al. (1991) investigates the effect of moderate heat stress on levels of cortisol in three groups of women: 1) non-pregnant women, (2) women 13-14 weeks pregnant, and (3) women 36-37 weeks pregnant. Blood samples were taken every 5–10 minutes during a resting period followed by the heat stress intervention. Although this study uses a relatively small sample, the authors found that exposure to moderate heat stress increased significantly cortisol levels in pregnant women. While this study does not examine offspring outcomes, it does suggest that changes in cortisol levels could be an important mediator. This evidence is important in view of the growing consensus that prenatal exposure to increased cortisol levels negatively impacts offspring outcomes. Indeed, a set of recent studies have found that exposure to high increased levels of cortisol is associated with impaired brain development at three and eight months, and with worse health outcomes (Aizer, Stroud, and Buka 2012; Huizink et al. 2003).

Considering this, it has been hypothesized that prenatal heat stress can have adverse consequences on child outcomes. Randomized studies based on animals have established a strong link between *in utero* exposure to extreme hot temperatures and poor offspring outcomes. These works generally exogenously expose pregnant animals to high temperatures. Examples include Shiota and Kayamura (1989) who exposed mice to high temperatures during pregnancy and observed retardation in brain growth of offspring. Strong, Silva, Cheng, and Eicher (2015) likewise exposed pregnant cows to hot temperatures and found that the offspring of exposed cows exhibited poorer health outcomes, including damages in immune system.

As these studies show that heat stress during the prenatal period results in diminished offspring outcomes, they are keys to extrapolate findings based on animal experiments to humans. This supports evidence from non-randomized studies that exposure to extreme hot temperatures during pregnancy is associated with poor infant outcomes. Much of this evidence becomes from epidemiological literature. In general, this literature finds that exposure to higher temperatures is associated with increased risk of prematurity and low birth weight.⁵ But these studies suffer from problems of endogeneity as they are based on cross-sectional or time series comparisons. Surprisingly, the most convincing evidence comes from works in economics. For

³ Some studies using twin-fixed effects have overcome this problem. However, those works based on comparisons of non-twin siblings must deal with the problems mentioned above.

⁴ There are other potential mechanisms through which extreme hot temperatures may affect initial endowments, including increases in disease transmission like malaria. It has been well documented that prenatal exposure to malaria is associated with poorer birth outcomes (e.g, Sarr et al. 2006). However, we believed that such mechanism is likely to be less prominent since the incidence of malaria is not high in Colombia. While this potential mechanism is not emphasized through article, we do not rule out the possibility that it plays a role.

⁵ A full review of the epidemiological literature can be found in Strand, Barnett, and Tong (2011).

instance, Deschenes, Greenstone, and Guryan (2009) exploit plausible exogenous variation in temperature within counties in U.S and find that prenatal heat stress during second trimester negatively affects birth-weight. Other studies using a similar approach find also negative impacts on APGAR scores and prematurity (Andalón et al. 2014). In general, this body of evidence suggests that prenatal heat stress could have adverse consequences on child endowments.

3. Data

Our analysis is based on children who are under five years of age with at least a sibling. Below, we describe the weather, investment, and supplementary data that we will analyze. Investment outcomes data are available for cohorts born 1990–2010. To identify exposure to heat waves during pregnancy, these data are matched to the weather measures based on the date of birth and the mother’s municipality of residence. Summary statistics of these data are presented in Table 1.

3.1. Weather Dataset

We have built a series for temperature and precipitation using data from the Terrestrial Air Temperature and Terrestrial Precipitation: 1900–2010 Gridded Monthly Time Series, version 3.02, respectively (Matsuura and Willmott 2012). This dataset provides worldwide estimates for weather conditions at the 0.5×0.5 degree latitude/longitude grid.⁶ Using an interpolation algorithm, Matsuura and Willmott (2012) computes values for each grid node from several nearby weather stations. Since some years did not have weather stations over the entire period, the data for missing years are imputed using a meteorological model. To minimize any potential bias from this measurement error, we focus on the period 1970–2010, as most of the weather stations were established in Colombia from 1970 and onwards. We use a strategy similar to Rocha and Soares (2015) to construct a municipality-by-month of weather panel. To start, we compute the centroid for each of the 1,120 municipalities in Colombia. Then, using the centroid, we located the four closest nodes to build a monthly series of temperature and precipitation as the weighted average of estimates related to these four nodes. As weights, we use the inverse of the distance to each node. The mean per municipality per month of temperature in our sample is 21.5 °C, with a standard deviation of 4.7 °C.

Using this consolidated dataset, we define a heat wave for a given month as temperature above the 90th percentile of distribution for that calendar month within the municipality. Since we are not comparing municipalities, the “extreme” hot temperature should not be taken in an absolute sense. These are simply extreme high temperature months for each municipality within the given time frame. We also investigate the effects of less severe heat waves by defining heat wave as temperature above 85th, 80th, and 75th percentiles.

Prenatal exposure to heat waves through pregnancy is measured by *first trimester*, *second trimester*, and *third trimester*. If, for example, a child was born on October, then *first trimester* is calculated as the number of extreme high temperature months that occurred in their municipality of birth during the months of February, March and April. Naturally, the *second trimester* is computed by the number of extreme high temperature months that occurred in their municipality of birth during the months of May, June, and July; and *third trimester* is computed using these criteria during the months of August, September, and October.

3.2. Main Outcomes

Our empirical analysis uses the 1995, 2000, 2005 and 2010 waves of the Demographic and Health Survey (DHS) of Colombia, a nationally representative survey of women ages 15 to 49. The DHS contains detailed

⁶ 0.5 degree correspond to 56 kilometers.

information on early-life health investments for all children under five. For our analysis, we pooled these DHS waves into one dataset. We restricted the sample to mothers with at least two children given that we used family fixed effects models. We also focus on children who were more than 12 months old at time of the survey. Our basic sample consists of 8,949 children with at least one sibling. We use the municipality of residence as a proxy for child's municipality of birth. This is a reasonable proxy given the low migration rates of infants.

Our health investments variables include vaccinations, and breastfeeding. Available measures of vaccination reported consistently across the four waves of the DHS include: polio, DPT (diphtheria, pertussis and tetanus combination), and measles. In Colombia, the recommend vaccination schedule is: polio at two months, four months, and six months; DPT at two months, four months, and six months; measles at 11 months. Our analysis investigates the effect of prenatal heat on the likelihood of being vaccinated for specific diseases. In terms of breastfeeding, we use a dummy variable that equals to one if the child was breastfed for more than six months. This is the minimum length of breastfeeding recommended by the World Health Organization. In the 2005 DHS, breastfeeding duration is only recorded for the youngest child born to a surveyed mother. Therefore, we exclude children from the 2005 DHS for the breastfeeding analysis.

3.3.Other Data

As a complementary analysis, we use the birth certificate microdata for the period 1998-2010 from the Colombian Department of Statistics (DANE). We obtained these administrative data for all the municipalities in Colombia- approximately 8 million birth records. This register provides date of delivery, information on gestation length, weight, and APGAR scores. In Colombia, there are between 400,000 and 700,000 births per year.⁷ Using this information, we construct a municipality-by-month of birth data set for the 1998-2010 period. The municipality of reference in this panel is that where the mother lives at birth. Our outcomes of interest are rate of low 5 minute APGAR (<8), rate of very low birth-weight ($\leq 1,500$ gr.), rate of low birth-weight ($\leq 2,500$ gr.), rate of birth via Caesarian section and rate of prematurity (less than 37 weeks of gestation).⁸

For further analysis, we also use microdata from the 2005 demographic census (the most recent available). The Integrated Public Use Microdata Series (IPUMS) provides a one percent sample. Although the census does not collected information on parental investments, we can assess whether selective migration may drive our main results. Using these data, we also estimate the long-term impacts of prenatal exposure to heat waves. The census does not have any measure of earnings, but there is information on educational attainment. Educational attainment is particularly interesting, as it is an important determinant of other later-life outcomes, including earnings, health and productivity (Cutler and Lleras-Muney 2010; Oreopoulos 2006). Thus, we investigate whether prenatal heat waves affect total years of schooling for adults aged 20-36.

Finally, we use other data sources for supplementary analysis. To assess whether our main results may be driven by changes in the local economic activity, we collected data on: *i*) municipality-year level information on local public revenue and spending collected by the Economics Research Center at Andes University for

⁷ Since there is no unique mother identifier in the data, subsequent births by the same mother cannot be identified. This precludes the use of the sibling-fixed effects estimator.

⁸ APGAR score is a clinical test that is given to the newborn in which five parameters are assessed. The parameters evaluated are muscle tone, respiratory effort, heart rate, reflexes and skin color. The test provides a total score between 0 and 10, where a higher score means healthier.

the period 1993-2010; ii) *departemento*-year level data on Gross Domestic Product (GDP) and Agricultural production (available for the period 1990-2010) from the DANE.⁹

3.4. Variation in Prenatal Heat Stress Within Families

An important concern about the sibling analysis is that siblings may experience “too similar” prenatal exposure to extreme hot temperatures. This may weaken the within-sibling relation between parental investments and prenatal heat stress. However, this argument has little empirical support. Prenatal exposure to heat waves varies widely across children in our sample. The standard deviation in the number of months exposed to hot temperatures during pregnancy is 1.58 (relative to a mean of 0.94). More importantly, mother fixed-effects explain only about 48 percent of the variation in the number of months exposed to heat waves while *in utero*, leaving a fair amount of within-sibling variation. This wide within-sibling variation is the basis of our identification strategy.

4. Empirical Strategy

Equation (1) relates each parental investment, y , of the child (i) born from mother (j) in municipality (k) to the three measures of prenatal heat waves discussed above:

$$y_{ijktm} = \alpha + \beta_1(1sttrimester)_{ijktm} + \beta_2(2ndtrimester)_{ijktm} + \beta_3(3rdtrimester)_{ijktm} + \delta'X_{ijktm} + \eta_j + \lambda_t + \mu_m + \xi_{ijktm} \quad (1)$$

where the vector X_{ijktm} includes indicators for child's gender, birth order, age in months, and municipal-rainfall for each trimester. λ_t , and μ_m are fixed effects for year of birth and month of birth, respectively. Because η_j is included on the right-hand side of the equation, it is used only within-sibling variation to identify the parameters β_1 through β_3 .

Model (1) essentially uses sibling differences in prenatal exposure to extreme hot temperatures, the timing of which is plausibly exogenous, give the unpredictability of a temperature shock, to identify prenatal heat stress impacts. Thus, our strategy compares parental investments of children prenatally exposed to greater extreme hot temperature months against the parental investments for siblings exposed to less extreme hot temperature months. The reason why one of them ended up with a greater exposure and the other one did not can be, for all practical purposes, considered as random. Given the evidence of previous studies that heat stress during pregnancy has negative consequences on initial endowments, positives values for β_1 , β_2 , and β_3 are interpreted as compensatory parental responses and negative values are interpreted as reinforcing responses.

Using this empirical approach, we are able to identify the causal impact of prenatal extreme hot temperature on parental responses. A potential problem pervading our analysis is related to recent evidence that parents may be changing fertility decisions based on temperature around the time of conception. Barreca, Deschenes, and Guldi (2015) convincingly show that parents are likely to postpone conception by one month in response to additional extreme high temperature. To the extent that this is important and that these parents may differ in ways that could affect parental inputs, between-family estimates of the effect of heat stress during first trimester may be biased. Our approach deals with this issue as it relies on within-family comparisons, thereby controlling any time-invariant family qualities. The use of this strategy would be biased if the specific-shift in the timing of conception is directly related to future family postnatal investments. There is no reason to believe that this is plausible. In addition, we provide a variety of evidence that our

⁹ *Departamento* is a first-order administrative unit similar to U.S States. In Colombia, there are 33 *departamento*.

results are unlikely to be driven by time series correlation in temperature, selective fertility, migration, or changes in local economic activity.

5. Results

5.1. Main Results

Our main results are presented in Table 2. Columns (1)-(4) look at vaccination during infancy. We use dummy variables indicating whether the child has the recommend vaccination doses for specific diseases. Column (4) uses a dummy variable that equals one if the child has all recommend vaccination doses. The results for breastfeeding are presented in columns (5).

The results from (1)-(4) show that exposure to heat waves during second trimester significantly increases the likelihood of being vaccinated. The magnitudes of the effects vary depending on the vaccination. One additional month of exposure to heat waves during the second trimester increases the probability of receiving the recommend vaccination schedule for polio, DPT, and measles by 2, 3, and 1.5 percentage points, respectively. In column (5), we find that exposure to prenatal heat stress during second trimester is significantly associated with increases in probability of having been breastfed for more than six months (point estimate of 0.03).

Overall, the evidence suggests that prenatal heat stress increases health investments. One way to assess the size of the effects is to compare them to the impacts of early-life interventions. For example, Attanasio et al. (2005) show that *Familias en Accion* (FA), a conditional cash transfer program in Colombia, increases the probability of DPT vaccination by 9 percentage points. This shows that the estimated effect of second trimester exposure on DPT vaccination is on third of the effect of the FA program. In other words, the child from a mother who was exposed to three extreme high temperature months during the second trimester would have a similar probability of receiving DPT vaccination as a child of a mother enrolled in the FA program.

5.2. Heterogeneity by severity of intensity

Our baseline specifications estimate the impacts of prenatal exposure to very extreme hot temperatures. A natural extension is to assess the presence of heterogeneous effects with respect to the severity of the shock. Table 3 explores this question by using measures of prenatal exposure that define heat waves as temperature above 85th, 80th, and 75th of distribution. In general, we find in fact that exposure to less extreme hot temperatures has significant smaller effects. For example, the marginal effect of exposure during second trimester ranges on the likelihood of total vaccination ranges from 3.5 percentage points in the baseline estimate to 1.6 percentage points in specification that uses the least extreme measure of heat wave. This analysis highlights the usefulness of an intensity-specific analysis when assessing the effects of environmental shocks *in utero*.

5.3. Potential Mechanisms and Robustness Checks

Next, we explore potential mechanism by which maternal heat stress affects parental investments. The results from this section suggest that variations in child endowments is a plausible explanation to our results. While the evidence is supportive of this idea, alternative interpretations may be also consistent with the patterns in parental investments. As we shall see, such alternative hypothesis have little empirical support.

5.3.1 Heat stress and child endowments

Studies based largely on animal experiments suggest that exogenous exposure to heat stress *in utero* negatively affects offspring endowments. This and discussions in economics that returns to child investments depend on initial endowments suggest that variations in birth endowments could be an important mediator between maternal heat stress and parental investments. While there is extensive literature documenting that prenatal heat stress has adverse consequences on fetal health (see, for example, Deschenes, Greenstone, and Guryan (2009)), we also test for the relationship using birth certificate data for the period 1998-2010. The results are presented in Table 4. We weight observations by the number of births per month in the municipality.

Panel A uses a specification that adjusts only for the baby's sex, for municipality-rainfall in each trimester and for municipality of residence at birth, year of birth and month of birth fixed effects. We find a positive and significant effect of prenatal exposure to heat stress during first trimester on very low birth weight. Increasing the number of extreme high temperature months by 1 increases the probability of very low birth weight by 0.009 percentage points (P -value=0.052). Compared to the mean of 0.8 percent, the effect is 1.1 percent. We also find that prenatal heat stress during second trimester has a positive and large effect on the likelihood of having a low APGAR score. One additional month exposed during second trimester increases the incidence of low 5 minute APGAR by 0.98 percentage points. Relative to the mean low 5 minute APGAR rate of 2.3 percent, the effect is substantial at 40 percent. Almond, Chay, and Lee (2005) show that 5 minute APGAR may be a more reliable measure of fetal health than birth weight.

The remaining columns examine the effect on other proxies for fetal health. In column (4) we show that an additional heat wave in the second trimester increases the probability of having a caesarian section by 0.16 percent points (p -value =0.04). We consider that this variable potentially reflects the presence of problems at birth, which may be correlated with an increased risk of poor infant health. The specific timing of the effect is consistent with Currie and Rossin-Slater (2013) who find that exposure to stress induced by an extreme weather event during the second trimester increases the probability of having a caesarian section in U.S. In column (5) we also find that exposure to prenatal heat in the third trimester leads to preterm birth.

Panel B corresponds to specifications that include maternal characteristics as control variables. The estimated coefficients are insensitive to adding such additional variables. In general, the estimates are significant and imply that prenatal exposure to heat waves is associated with poorer health at birth. Furthermore, in many cases, their precision improves. This provides reassuring evidence on the validity of the empirical approach.

Overall, the findings support the notion that heat waves can have adverse consequences on the infant health, confirming the evidence from prior studies. It is important to recognize that this analysis may underestimate the impact of heat stress, as we do not use other more direct measures of infant health. For example, Currie and Rossin-Slater (2013) show convincingly that prenatal stress induced by extreme weather events has a substantial larger effect on the probability of complications of labor and delivery, and of abnormal conditions such as meconium aspiration syndrome. In view of this evidence and that these variables could be more direct proxies for health at birth, our estimates can be interpreted as lower bounds of the effect of maternal heat stress on infant health.

5.3.2 Migration

Given that we use the municipality of residence as a proxy for child's municipality of birth, a bias could drive our results if municipality-migration is related to extreme hot temperatures. It is hard to argue that this is the case since we have used temporary variations in temperature and they are unable to cause disruption of physical infrastructure (unlike others weather events such as storms). To asses this more formally, we have

examined the 2005 census and analyzed differences in prenatal heat waves between migrant and non-migrant children. In Table 5 we regress prenatal exposure to heat waves on a dummy variable that equals 1 if the child was born in the survey municipality. Consistent with the view that heat waves are unlikely to be related to migration, we find no differences in prenatal exposure to extreme hot temperatures between migrant and non-migrant children.

Still, the use of municipality of residence as a proxy for municipality of birth most likely introduces a random measurement error that attenuates our estimates. To investigate the magnitude of this potential bias, in Table 6 we estimate the investment regressions based on a sample that includes only mothers with children who were born in the municipality of residence. The results are qualitatively similar to our baseline estimates. While in general the estimated coefficients of second trimester exposure are larger in magnitude, they are very similar to those of the baseline estimates. Collectively, these findings are consistent with the presence of a random measurement error and suggest that the resulting attenuation bias in our sample is small.

5.3.3. *Subsequent fertility*

Some studies argue that child endowments can affect future fertility decisions. This would be a potential channel explaining the link between heat stress, initial endowments and parental investments. The seminal study by Becker and Tomes (1976) suggests ambiguous predictions on the direction of how variations in child endowments may affect future fertility. On the one hand, if less healthy children increase the cost of child quantity, then it would lead to a reduction in fertility. Alternatively, if returns to child quality are lower in less endowed children, then this would increase fertility in response to the higher shadow price of child quality. However, we are unable to find any significant effect of heat waves on fertility. We do this in Table 7 by checking whether prenatal heat stress affects the quantity and spacing of children. The estimates are small and tightly bound around zero, suggesting that having a child who was prenatally exposed to heat waves did not alter subsequent fertility decisions significantly.

5.3.4. *Other hypothesis*

Perhaps an obvious objection to our interpretation of the main results comes from recent evidence that high temperatures shocks lead to economic downturns (Dell, Jones, and Olken 2012). Therefore, one may argue that the patterns in parental investments are in fact not so much determined by child endowment shifts, but by parents' reactions to reductions in the cost of opportunity, given that the inputs we have used are time-intensive. This hypothesis is somewhat consistent with Miller and Urdinola (2010) who show that time-intensive investments are higher during economic downturns. Yet, for this alternative interpretation to make sense, agricultural sector should be an important sector in the Colombian economy. This does not seem to be the case. Indeed, the agricultural value-added only accounts for 11 percent of GDP across the 1990-2010 period, which contrasts with countries like Uganda where agriculture importance ranges from 30 to 60 percent of GDP. Furthermore, the available evidence indicates that higher temperature leads to economic downturns only in poor countries. Even if temperature shock lead to economic downturn, it is hard to think of reasons why this would explain significant effects on parental investments in urban children since such economic shocks would be presumably important in rural areas. In fact, we find little evidence of differential impacts between rural and urban children.¹⁰

In any case, we can directly investigate this alternative hypothesis by estimating the relationship between heat waves and production, using *departemento*-year level data on GDP and agricultural production

¹⁰ These analyses are available upon author request.

for the period 1990-2010, and municipality-level data on local public finance for years 1993 through 2010. We use data on public finance as proxies for local economic activity. The variable independents of interest are the number of extreme high temperature months that occurred in the *departamento*/municipality in a given year, along with a one-year lag. The results are presented in Table 8, with dependent variables shown in the first column of each row. Considering the discussion above, it is not surprising to find insignificant estimates on these regressions. Moreover, the estimated coefficients are very small in magnitude. We take these results as evidence that our main findings are in fact not driven by changes in local economic activity.

Other alternative hypothesis is that heat waves this month may be correlated with heat waves next months. If so, our estimates may not represent the effects of heat waves while *in utero*. While this is a reasonable hypothesis, we can test it by including in the same regressions the variables of heat exposure after birth. We do this in Table 9. The coefficients associated with heat stress *in utero* during second trimester remain virtually identical, casting doubt on this alternative explanation.

One could argue that was not just child endowment shifts that affect parental investments, but also reductions in time allocated to labor to minimize the potential health impacts of warmer temperatures. This idea is made more plausible in view of evidence from Graff Zivin and Neidell (2014), who find a moderate short-run (within few weeks after the shock) decline in time allocated to labor at high temperature. This alternative hypothesis may explain why parents increase health investments, as vaccination and breastfeeding inputs are made early in life. However, the evidence from Table 2 that exposure during the third trimester does not affect parental investments weights against this alternative interpretation, as one would expected to observe an effect of exposure around the time of birth. Note also that, to the extent that this explanation is important, we should see changes in our estimates when controlling for heat waves after birth. We do not. Although we cannot completely rule out this possibility, the evidence suggests that it is unlikely to be the main mechanism driving the patterns in parental investments.

5.3.5. Further robustness checks

We conduct additional robustness checks.¹¹ First, we examine the relationship between prenatal heat stress and parental investments by using specifications that control for municipality-specific linear time trends. In general, the use of this more demanding specification produces estimated coefficients of the second trimester exposure that are very similar and that remain statistically significant. Second, we exclude twin children given that prenatal exposure between them does not vary. Our results are broadly similar. Finally, we also assess whether there are heterogeneous effects by child's sex, rural/urban location, and mother's education. We then present estimates that interact second trimester exposure with these characteristics. We find differential impacts that are statistically significant only in a few cases. For instance, the compensating investment behavior in terms of breastfeeding is more pronounced in boys than girls. Still, there are no significant differences between boys and girls in vaccinations. We conclude that there is little evidence of a consistent interaction between prenatal heat exposure and these characteristics.

6. Long-Term Effects on Human Capital Accumulation

We also explore whether there are long-term impacts of exposure to extreme hot temperatures on the accumulation of human capital. If parental health inputs are important enough in the production of child quality, then the prenatal temperature shock may be completely remedied. To explore this question, we use data from 2005 census and estimate a specification similar to Eq. (1), but in this case we use total years of schooling for cohorts born 1970–1986 (ages 20-36 at time of survey) as a dependent variable. The census

¹¹ Given space constraints, we do not present the results for these robustness exercises, but they are available upon author request.

only keeps information on the sibling composition of those individuals who co-reside with their relatives. As adults are less likely to live with their parents than children, we do not include sibling-fixed effects. In practice, we include a robust set of municipality of birth, year of birth and month of birth fixed effects.

In Table 10, the results indicate that exposure during second trimester of pregnancy to heat waves is associated with lower levels of schooling. The effect in magnitude is small (0.3 percent relative to the mean). Column (3) uses a specification that interacts second trimester with female. While the effect appears to be stronger for men, this heterogeneity is not statistically significant. The results from columns (4)-(6) show that the estimated coefficient is robust to controlling for exposure to heat waves after birth. Overall, in this section, we provide evidence that the effects of prenatal heat waves can last into adulthood, suggesting that compensatory health investments during infancy are not sufficient to remedy the initial health insult.

7. Conclusion

In this paper we estimate the impacts of *in utero* exposure to heat waves on parental investments using Colombian data. We find that prenatal exposure to heat waves is associated with more postnatal health investments. A variety of evidence is presented in favor of the interpretation that this parental behavior reflects compensating responses to shifts in children's endowments. We interpret these findings as evidence highly consistent with the model of intrahousehold resource allocation from Yi et al (2015). In particular a possible interpretation is that under substitutability between health and investment in health, endowment shifts induced by prenatal heat stress would increase the returns to child health quality. As a result, parents responded by devoting more resources in health.

These results relate to a growing literature in economics that documents long-term effects of exposure *in utero* to environmental shocks. The findings in this paper are consistent with previous studies indicating that family investments respond to child endowments. One implication is that reduced-form estimates of the effect of environmental conditions *in utero* on later life outcomes do not necessarily represent a biological effect. Therefore, caution is necessary when interpret estimates from such studies.

Our findings have important policy implications. Climatic projections indicate that the climate's earth will become hotter in the coming century, which would make more frequent the incidence of extreme weather events. The long-term effects of such extreme weather events on human capital accumulation several decades later should be factored into cost-benefit analyses of climate change mitigation policies. The evidence that the initial shock is not completely remedied by parental investments provides additional justification for interventions that shield pregnant women from the consequences of temporary environmental shocks.

References

- Adhvaryu, Achyuta, and Anant Nyshadham. 2014. "Endowments at Birth and Parents' Investments in Children." *The Economic Journal*.
- Aizer, Ana, L Stroud, and S Buka. 2012. *Maternal Stress and Child Outcomes: Evidence from Siblings*. NBER Working Papers 18422.
- Aizer, Anna, and Flávio Cunha. 2012. *The Production of Human Capital: Endowments, Investments and Fertility*. NBER Working Paper 18429.
- Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post 1940 U.S. Population." *Journal of Political Economy* 114 (4): 672–712.

- Almond, Douglas, Kenneth Y Chay, and David S Lee. 2005. "The Costs of Low Birth Weight." *The Quarterly Journal of Economics* 120 (3): 1031–83.
- Almond, Douglas, and Bhashkar Mazumder. 2013. "Fetal Origins and Parental Responses." *Annual Review of Economics* 5 (1). Annual Reviews: 37–56.
- Andalón, Mabel, João Pedro Azevedo, Carlos Rodríguez-Castelán, Viviane Sanfelice, and Daniel Valderrama. 2014. *Weather Shocks and Health at Birth in Colombia*. Policy Research Working Paper 7081.
- Anderson, J. W., B. M. Johnstone, and D. T. Remley. 1999. "Breast-Feeding and Cognitive Development: A Meta-Analysis." *American Journal of Clinical Nutrition* 70 (4): 525–35.
- Attanasio, Orazio, Carlos Gómez, Patricia Heredia, and Marcos Vera-Hernández. 2005. *The Short-Term Impact of a Conditional Cash Subsidy on Child Health and Nutrition in Colombia*. London: Institute for Fiscal Studies
- Barreca, Alan, Olivier Deschenes, and Melanie Guldi. 2015. *Maybe Next Month? Temperature Shocks, Climate Change, and Dynamic Adjustments in Birth Rates*. NBER Working Paper 21681.
- Becker, Gary S, and Nigel Tomes. 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy* 84 (4): 143–62.
- Behrman, Jere R, Robert A Pollak, and Paul Taubman. 1982. "Parental Preferences and Provision for Progeny." *Journal of Political Economy* 90 (1): 52–73.
- Behrman, Jere R, and Mark R Rosenzweig. 2004. "Returns to Birthweight." *Review of Economics and Statistics* 86 (2): 586–601.
- Currie, Janet, and Rosemary Hyson. 1999. "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight." *The American Economic Review* 89 (2): 245–50.
- Currie, Janet, and Maya Rossin-Slater. 2013. "Weathering the Storm: Hurricanes and Birth Outcomes." *Journal of Health Economics* 32 (3): 487–503.
- Cutler, David M, and Adriana Lleras-Muney. 2010. "Understanding Differences in Health Behaviors by Education." *Journal of Health Economics* 29 (1): 1–28.
- Datar, Ashlesha, M.Rebecca Kilburn, and DavidS. Loughran. 2010. "Endowments and Parental Investments in Infancy and Early Childhood." *Demography* 47 (1): 145–62.
- Davis, Poggi, and Curt A. Sandman. 2010. "The Timing of Prenatal Exposure to Maternal Cortisol and Psychosocial Stress Is Associated with Human Infant Cognitive Development." *Child Development* 81 (1): 131–48.
- Del Bono, Emilia, John Ermisch, and Marco Francesconi. 2012. "Intrafamily Resource Allocations: A Dynamic Structural Model of Birth Weight." *Journal of Labor Economics* 30 (3): 657–706.

- Del Bono, E., and B. Rabe. 2012. *Breastfeeding and Child Cognitive Outcomes: Evidence from a Hospital-Based Breastfeeding Support Policy*. ISER Working Paper Series 2012-29, University of Essex.
- Dell, M, B Jones, and B Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." *American Economic Journal: Macroeconomics* 4: 66–95.
- Deschenes, Olivier, Michael Greenstone, and Jonathan Guryan. 2009. "Climate Change and Birth Weight." *American Economic Review* 99 (2): 211–17.
- Deschênes, Olivier, and Enrico Moretti. 2009. "Extreme Weather Events, Mortality, and Migration." *Review of Economics and Statistics* 91 (4): 659–81.
- Graff Zivin, Joshua, and Matthew Neidell. 2014. "Temperature and the Allocation of Time: Implications for Climate Change." *Journal of Labor Economics* 32 (1): 1–26.
- Huizink, Anja C, Pascale G Robles de Medina, Eduard J H Mulder, Gerard H A Visser, and Jan K Buitelaar. 2003. "Stress during Pregnancy Is Associated with Developmental Outcome in Infancy." *Journal of Child Psychology and Psychiatry* 44 (6): 810–18.
- IPCC. 2007. *Fourth Assessment Report: Climate Change 2007. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland.
- Kelly, Elaine. 2011. "The Scourge of Asian Flu: In Utero Exposure to Pandemic Influenza and the Development of a Cohort of British Children ." *Journal of Human Resources* 46 (4): 669–94.
- Matsuura, K, and C. Willmott. 2012. *Terrestrial Air Temperature and Precipitation: 1900-2010 Gridded Monthly Time Series, Version 3.02*.
- Miller, Grant, and B Piedad Urdinola. 2010. "Cyclicalities, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Colombia." *The Journal of Political Economy* 118 (1): 113–55.
- Oreopoulos, Philip. 2006. "Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter." *The American Economic Review* 96 (1): 152–75.
- Parman, John. 2013. *Childhood Health and Sibling Outcomes: The Shared Burden and Benefit of the 1918 Influenza Pandemic*. NBER Working Paper 19505.
- Rocha, Rudi, and Rodrigo R Soares. 2015. "Water Scarcity and Birth Outcomes in the Brazilian Semi-arid." *Journal of Development Economics* 112 (0): 72–91.
- Rosenzweig, Mark R, and Kenneth I Wolpin. 1988. "Heterogeneity, Intrafamily Distribution, and Child Health." *The Journal of Human Resources* 23 (4): 437–61.
- Sarr, Demba, Laurence Marrama, Alioune Gaye, Jean M Dangou, Makhtar Niang, Odile Mercereau-Puijalon, Jean Y Lehesran, and Ronan Jambou. 2006. "High Prevalence of Placental Malaria and Low Birth

- Weight in Sahelian Periurban Area.” *The American Journal of Tropical Medicine and Hygiene* 75 (1): 171–77.
- Shiota, K, and T Kayamura. 1989. “Effects of Prenatal Heat Stress on Postnatal Growth, Behavior and Learning Capacity in Mice.” *Neonatology* 56 (1): 6–14.
- Strand, Linn B, Adrian G Barnett, and Shilu Tong. 2011. “The Influence of Season and Ambient Temperature on Birth Outcomes: A Review of the Epidemiological Literature.” *Environmental Research* 111 (3): 451–62.
- Strong, R A, E B Silva, H W Cheng, and S D Eicher. 2015. “Acute Brief Heat Stress in Late Gestation Alters Neonatal Calf Innate Immune Functions.” *Journal of Dairy Science* 98 (11): 7771–83.
- Totsika, Vasiliki, and Kathy Sylva. 2004. “The Home Observation for Measurement of the Environment Revisited.” *Child and Adolescent Mental Health* 9 (1): 25–35.
- Vaha-Eskeli, Kalevi, Risto Erkkola, Kerttu Irjala, and Ormo Viinamaki. 1991. “Effect of Thermal Stress on Serum Prolactin, Cortisol and Plasma Arginine Vasopressin Concentration in the Pregnant and Non-Pregnant State.” *European Journal of Obstetrics and Gynecology and Reproductive Biology* 42: 1–8.
- Venkataramani, Atheendar S. 2012. “Early Life Exposure to Malaria and Cognition in Adulthood: Evidence from Mexico.” *Journal of Health Economics* 31 (5): 767–80.
- Victora, C. G., M. de Onis, P. C. Hallal, M. Blössner, and R. Shrimpton. 2010. “Worldwide Timing of Growth Faltering: Revisiting Implications for Interventions.” *Pediatrics* 125 (3): e473–80.
- Wadhwa, Pathik D, Curt A Sandman, Manuel Porto, Christine Dunkel-Schetter, and Thomas J Garite. 1993. “The Association between Prenatal Stress and Infant Birth Weight and Gestational Age at Birth: A Prospective Investigation.” *American Journal of Obstetrics and Gynecology* 169 (4): 858–65.
- Wells, Jonathan C K, and Tim J Cole. 2002. “Birth Weight and Environmental Heat Load: A between-Population Analysis.” *American Journal of Physical Anthropology* 119 (3): 276–82.
- Wilde, Joshua, Bénédicte Apouey, and Toni Jung. “Heat Waves at Conception and Later Life Outcomes.” Forthcoming in *European Economic Review*.
- Yi, Junjian, James J Heckman, Junsen Zhang, and Gabriella Conti. 2015. “Early Health Shocks, Intra-Household Resource Allocation and Child Outcomes.” *The Economic Journal* 125 (588): 347–71.

Tables

Table 1. Descriptive statistics, Demography Health Survey (1995, 2000, 2005 and 2010 rounds)

	Mean	Standard deviation
Complete schedule of vaccination (Polio, DPT and Measles doses)	0.60	0.48
Complete schedule of Polio vaccination	0.72	0.44
Complete schedule of DPT vaccination	0.80	0.40
Measles vaccination	0.83	0.37
Breastfed for more than six months	0.68	0.46
Child is male	0.51	0.50
Child's age in months	35.50	1.45
First born	0.21	0.41

Note. Survey weights are used.

Table 2. Effects of Heat Waves *In Utero* on Parental Investments

	Health investments				
	Vaccinations				Breastfed for more than six months
	Polio	DPT	Measles	Total vaccination	
	(1)	(2)	(3)	(4)	(5)
<i>1st Trimester</i>	-0.000 [0.011]	-0.005 [0.011]	-0.010 [0.008]	-0.001 [0.011]	0.006 [0.016]
<i>2nd Trimester</i>	0.022 [0.009]**	0.029 [0.010]***	0.015 [0.009]*	0.035 [0.012]***	0.031 [0.015]**
<i>3rd Trimester</i>	0.014 [0.010]	-0.003 [0.008]	-0.008 [0.009]	-0.004 [0.009]	0.002 [0.015]
N	8,949	8,949	8,949	8,949	5,932

Notes. Robust standard errors in brackets (***p < 0.01, **p < 0.05, *p < 0.1) are clustered at the municipality level. All regressions are based on the mother-effects fixed estimator, and include as controls dummies for age in months, precipitation in each trimester, dummies for year of birth, dummies for month of birth, and child's sex and birth order.

Table 3. Effects of Heat Waves *In Utero* on Parental Investments (Heterogeneity by severity of intensity)

	Temperature above:			
	90 th percentile (1)	85 th percentile (2)	80 th percentile (3)	75 th percentile (4)
<i>Panel A: Dependent Variable is Polio vaccination</i>				
2nd Trimester heat exposure	0.022 [0.009]**	0.025 [0.008]***	0.016 [0.007]**	0.019 [0.007]**
N	8,949	8,949	8,949	8,949
<i>Panel B: Dependent Variable is DPT vaccination</i>				
2nd Trimester heat exposure	0.029 [0.010]***	0.018 [0.008]**	0.015 [0.006]**	0.010 [0.005]*
N	8,949	8,949	8,949	8,949
<i>Panel C: Dependent Variable is Measles vaccination</i>				
2nd Trimester heat exposure	0.015 [0.009]*	0.014 [0.007]**	0.009 [0.006]	0.005 [0.006]
N	8,949	8,949	8,949	8,949
<i>Panel D: Dependent Variable is Total vaccinations</i>				
2nd Trimester heat exposure	0.035 [0.012]***	0.027 [0.011]**	0.019 [0.009]**	0.016 [0.007]**
N	8,949	8,949	8,949	8,949
<i>Panel E: Dependent Variable is Breastfed for more than six months</i>				
2nd Trimester heat exposure	0.031 [0.015]**	0.017 [0.012]	0.009 [0.010]	0.005 [0.009]
N	5,932	5,932	5,932	5,932

Notes. Robust standard errors in brackets (***p < 0.01, **p < 0.05, *p < 0.1) are clustered at the municipality level. Column (1) produces baseline results. Columns (2) through (4) report estimates by using alternative definitions of prenatal heat waves.

Table 4. Effects of Heat Waves *In Utero* on Birth Outcomes

	Rate of very low birth weight (1)	Rate of low birth weight (2)	Rate of low 5 minute APGAR (3)	Rate of C-Section (4)	Rate of preterm birth (5)
<i>Panel A: Baseline specification</i>					
1st Trimester in utero	0.0099 [0.0050]**	0.0004 [0.0258]	-0.200 [0.441]	0.1161 [0.1271]	0.0036 [0.0405]
2nd Trimester in utero	-0.0060 [0.0044]	-0.0008 [0.0201]	0.9880 [0.5770]*	0.1662 [0.0839]**	-0.0776 [0.0589]
3rd Trimester in utero	0.0010 [0.0046]	0.0131 [0.0223]	0.3370 [0.3120]	0.03762 [0.1140]	0.0580 [0.0332]*
<i>Panel B: Controlling for maternal characteristics</i>					
1st Trimester in utero	0.0099 [0.0049]**	-0.0009 [0.0255]	-0.2050 [0.4461]	0.1321 [0.1272]	0.0044 [0.0401]
2nd Trimester in utero	-0.0061 [0.0045]	-0.0024 [0.0196]	0.9890 [0.5781]*	0.1711 [0.0836]**	-0.0792 [0.0591]
3rd Trimester in utero	0.0010 [0.0046]	0.0117 [0.0218]	0.3311 [0.3091]	0.0508 [0.1161]	0.0585 [0.0328]*
Mean of dependent variable	0.82	7.31	2.32	30.23	14.15
N	168,692	168,692	167,800	169,113	168,743

Notes. Robust standard errors in brackets (***p < 0.01, **p < 0.05, *p < 0.1) are clustered at the municipality level. Panel A is a specification that controls for municipality of birth, year of birth, and month of birth fixed effects, and for precipitation in each

trimester, and baby's sex. In addition, Panel B includes dummy for mother age under 20 years, dummy for mother age over 45 years, dummy for mother's education (some college), dummy for mother's marital status (married).

Table 5. Heat Waves *In Utero* between migrant and non-migrant children

	Number of prenatal heat waves during:		
	1st trimester (1)	2nd trimester (2)	3rd trimester (3)
Child was born in the survey municipality	0.021 [0.0158]	0.001 [0.014]	-0.012 [0.016]
N	1,222,311	1,222,311	1,222,311

Notes. Robust standard errors in brackets (**p < 0.01, *p < 0.05, *p < 0.1) are clustered at the municipality level. All regressions are based on the mother-effects fixed estimator, and include controls for precipitation in each trimester, dummies for municipality of birth, dummies for year of birth, dummies for month of birth, and child's sex and age.

Table 6. Effects of Heat Waves *In Utero* on Parental Investments (Excluding migrant families)

	Health investments				
	Vaccinations				Breastfed for more than six months (5)
	Polio (1)	DPT (2)	Measles (3)	Total Vaccination (4)	
<i>Panel A: Baseline estimates</i>					
<i>1st Trimester</i>	-0.000 [0.011]	-0.005 [0.011]	-0.010 [0.008]	0.001 [0.013]	0.006 [0.016]
<i>2nd Trimester</i>	0.022 [0.010]**	0.029 [0.010]***	0.015 [0.009]*	0.035 [0.012]***	0.031 [0.015]**
<i>3rd Trimester</i>	0.014 [0.010]	-0.003 [0.007]	-0.008 [0.009]	-0.004 [0.009]	-0.002 [0.015]
N	8,949	8,949	8,949	8,949	5,932
<i>Panel B: Excluding migrant families</i>					
<i>1st Trimester</i>	0.000 [0.012]	-0.008 [0.011]	-0.006 [0.008]	0.003 [0.013]	0.010 [0.018]
<i>2nd Trimester</i>	0.029 [0.009]***	0.039 [0.010]***	0.014 [0.010]	0.043 [0.014]***	0.040 [0.015]***
<i>3rd Trimester</i>	0.018 [0.012]	-0.009 [0.009]	-0.008 [0.011]	-0.005 [0.011]	-0.008 [0.014]
N	7,150	7,150	7,150	7,150	4,714

Notes. Robust standard errors in brackets (**p < 0.01, *p < 0.05, *p < 0.1) are clustered at the municipality level. Panel A produces baseline results. Panel B presents results based on a sample of children who were born in the municipality of residence.

Table 7. Effects of Heat Waves *In Utero* on Quantity and Spacing of Births

	Subsequent births (1)	Succeeding birth interval (2)
<i>1st Trimester</i>	-0.000 [0.004]	-0.600 [0.366]
<i>2nd Trimester</i>	0.001 [0.005]	0.335 [0.39]
<i>3rd Trimester</i>	0.002 [0.005]	0.484 [0.360]
N	8,949	4,923

Notes. Robust standard errors in brackets (**p < 0.01, *p < 0.05, *p < 0.1) are clustered at the municipality level. All regressions are based on the mother-effects fixed estimator, and include controls for mother's age at birth, precipitation in each trimester, dummies for year of birth, dummies for month of birth, and child's sex and birth order.

Table 8. Relationship between Heat Waves and Economic Activity

	(Heat waves) _t		(Heat waves) _{t-1}		N
	Coefficient Estimate	Standard Error	Coefficient Estimate	Standard Error	
Agricultural production	-0.001	[0.010]	0.005	[0.008]	640
GDP	0.005	[0.004]	0.004	[0.003]	660
Total Income (Indirect+ direct taxes)	-0.002	[0.004]	-0.003	[0.003]	17,257
Capital Income	-0.005	[0.006]	0.002	[0.006]	16,304
Transfer	0.005	[0.004]	0.000	[0.007]	16,791
Total Spending	-0.001	[0.004]	-0.003	[0.004]	17,316
Investment Spending	-0.004	[0.005]	0.002	[0.007]	11,676
Operational Spending	0.000	[0.002]	-0.007	[0.005]	17,236
Housing Spending	-0.001	[0.017]	-0.010	[0.020]	10,211
Education Spending	-0.008	[0.012]	0.000	[0.010]	11,662
Health Spending	-0.001	[0.007]	0.006	[0.007]	11,648

Notes. Robust standard errors in brackets (**p < 0.01, *p < 0.05, *p < 0.1) are clustered at the *departamento* level. Coefficients in each row are from a different regression. Regressions for agricultural production and GDP use data at *departamento* level and include *departamento* and year fixed effects, and control for precipitation and *departamento*-specific time trends. Regressions on local public finance use data at municipality level and include municipality and year fixed effects, and controls for precipitation and *departamento*-specific time trends.

Table 9. Effects of Heat Waves *In Utero* on Parental Investments (controlling for heat waves after birth)

		Controlling for heat waves:				
	Baseline	1 trimester after birth	2 trimester after birth	3 trimester after birth	4 trimester after birth	5 trimester after birth
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dependent Variable is Polio vaccination						
2nd Trimester heat exposure	0.022 [0.009]**	0.022 [0.009]**	0.024 [0.009]***	0.022 [0.009]**	0.022 [0.009]**	0.022 [0.009]**
Panel B: Dependent Variable is DPT vaccination						
2nd Trimester heat exposure	0.029 [0.010]***	0.029 [0.010]***	0.029 [0.010]***	0.028 [0.010]***	0.028 [0.010]***	0.028 [0.010]***
Panel C: Dependent Variable is Measles vaccination						
2nd Trimester heat exposure	0.015 [0.009]*	0.015 [0.009]*	0.016 [0.010]*	0.016 [0.010*]	0.017 [0.010]*	0.016 [0.009]*
Panel D: Dependent Variable is Total vaccinations						
2nd Trimester heat exposure	0.035 [0.012]***	0.035 [0.012]***	0.036 [0.012]***	0.034 [0.012]***	0.034 [0.012]***	0.034 [0.012]***
Panel E: Dependent Variable is Breastfed for more than six months						
2nd Trimester heat exposure	0.031 [0.015]**	0.031 [0.015]**	0.031 [0.014]**	0.028 [0.014]*	0.028 [0.014]*	0.028 [0.014]*

Notes. Robust standard errors in brackets (**p < 0.01, *p < 0.05, *p < 0.1) are clustered at the municipality level. Column (1) produces baseline results. Columns (2) through (6) report estimates controlling for heat waves after birth.

Table 10. Effects of Heat Waves *In Utero* on Total Schooling

	(1)	(2)	(3)	(4)	(5)	(6)
1st Trimester	0.007 [0.013]			0.007 [0.013]	0.005 [0.014]	0.005 [0.014]
2nd Trimester	-0.030 [0.014]**	-0.036 [0.014]**	-0.053 [0.015]***	-0.030 [0.014]**	-0.031 [0.014]**	-0.033 [0.014]**
3rd Trimester	-0.026 [0.030]			-0.022 [0.029]	-0.020 [0.029]	-0.021 [0.029]
2nd Trimester x Female			0.033 [0.021]			
1st Trimester after birth				-0.014 [0.017]	-0.008 [0.017]	-0.006 [0.017]
2nd Trimester after birth					-0.023 [0.022]	-0.016 [0.020]
3rd Trimester after birth						-0.025 [0.016]
N	9,092,354	9,092,354	9,092,354	9,092,354	9,092,354	9,092,354

Notes. Robust standard errors in brackets (**p < 0.01, *p < 0.05, *p < 0.1) are clustered at the municipality level. Dependent variable is years of education (Ages 20-36). The regressions contain municipality of birth, year of birth and month of birth fixed effects, and controls for age and sex.

New Evidence of the Effect of Body Weight on Labor Market Outcomes in Brazil

Carlos A. Charris¹ Bladimir Carrillo² Pedro Vertino³ Alexandre B. Coelho⁴

Resumo

Este artigo fornece novas evidências do efeito do peso corporal sobre os sucessos no mercado de trabalho no Brasil. A metodologia usa mudanças no peso corporal devidas a fatores genéticos como uma fonte de variação exógena. Os resultados mostraram que o peso corporal tem um efeito positivo sobre os salários e a probabilidade de emprego formal. Além disso, evidenciaram-se impactos heterogêneos por sexo, raça e área de residência. A evidência sugere também que a educação é um importante mecanismo através do qual o peso corporal afeta os sucessos no mercado de trabalho.

Palavras chaves: Peso corporal, sucessos no mercado de trabalho, emprego formal.

Classificação JEL: I10, J30, J24

Abstract

This paper provides new evidence of the effect of body weight on labor market outcomes in Brazil. The methodology exploits changes in body mass index due to genetic factors as a source of exogenous variation. We found that body weight has a positive effect on wages and formal employment. Furthermore, we also find heterogeneous effects by sex, race, and residence, and area. The evidence also suggests that education is an important mechanism through which body weight affects labor market outcomes.

Keywords: Body weight, Labor market outcomes, Formal employment.

JEL classification: I10, J24, J30

¹ Departamento de Economia Rural, Universidade Federal de Viçosa, Avenida Purdue, s/nº, Campus Universitário, Edifício Edson Potsch Magalhães 36570.900, Viçosa, MG, Brasil. E-mail: carlos.andres@ufv.br

² Departamento de Economia Rural, Universidade Federal de Viçosa, Avenida Purdue, s/nº, Campus Universitário, Edifício Edson Potsch Magalhães 36570.900, Viçosa, MG, Brasil. E-mail: bladimir.bermudez@ufv.br

³ Departamento de Economia Rural, Universidade Federal de Viçosa, Avenida Purdue, s/nº, Campus Universitário, Edifício Edson Potsch Magalhães 36570.900, Viçosa, MG, Brasil. E-mail: pedrovertino@hotmail.com

⁴ Departamento de Economia Rural, Universidade Federal de Viçosa, Avenida Purdue, s/nº, Campus Universitário, Edifício Edson Potsch Magalhães 36570.900, Viçosa, MG, Brasil. E-mail: acoelho@ufv.br

1. Introduction

This paper provides new evidence of the effect of body weight on labor market outcomes in a developing country context. Besides estimating the effects on wage, we also examine informal employment, one dimension of labor market that is prevalent in developing countries and has not previously been examined by literature. The literature has long discussed the relationship between body size, measured by body mass index (BMI), and labor market outcomes (BAUM; FORD, 2004; CAWLEY; HAN; NORTON, 2009; CAWLEY, 2004). It has been argued that obesity increases the risk of various diseases such as cancer, diabetes, depression and arthritis (ABBOTT et al., 1994; PI-SUNYER, 2002), adversely affecting labor productivity. This has been one of the main arguments in the literature to expect a negative relationship between BMI and wages, and much of the evidence, which has focused on the developed countries, is consistent with this view. However, such relationship may be different in developing countries. To the extent that the insufficient intake of nutrients is much more prevalent in developing countries, weight gains may reflect better health and nutrition, which could result in higher labor productivity. There is evidence consistent with this. Indeed, while increases on BMI lower wages in obese people, increases in BMI are associated with higher wages in non-obese individuals (SHIMOKAWA, 2008; WADA; TEKIN, 2010). This suggests that increases in BMI, on average, may lead to better labor market outcomes in developing countries.

Understanding the effects of body size on labor market outcomes is important for a number of reasons. First, an estimation of this relationship may provide relevant information on the determinants of poverty persistence due to that success in the labor market is directly related to family income and individual welfare. Second, understanding the effects of variations in the health and nutrition on labor productivity might contribute to formulation of employment benefit policies such as health insurance, health promotion programs and disease management interventions. Third, because body weight is related to health, the economic effects of events in the environment, such as environment-related disease, affecting health may be better quantified. This is particularly important in developing countries since adverse health shocks are more frequent and harder to be mitigated (CURRIE; VOGL, 2013).

Any empirical exercise that attempt to ascertain the importance of body weight in labor market outcomes must take into consideration the endogeneity of BMI. A source of endogeneity is simultaneity or reverse causality. Individuals with higher earnings may differently invest in their health. Moreover, if relevant variables, such as individual background, that are correlated with both labor productivity and BMI are omitted, the regressions will be subject to omitted variables bias, or the second source of endogeneity. Furthermore, if the anthropometric variables used to calculate the BMI are subject to measurement error, then this would lead to attenuation in the estimates. These sources of bias have been well discussed in the literature and various methods have been proposed to remedy. For developing countries, much of the existing studies have used disease environment, health infrastructure and health input prices as instruments for BMI. In the Brazilian case, Thomas and Strauss (1997) use the relative prices of food as instrumental variables for BMI and find that higher BMI leads to higher wages. The problem with using these instruments is that they could be correlated with other local characteristics (e.g. input supplies for firms) that may affect wages regardless of BMI.

Using Brazilian data, we explore the strategy proposed by Cawley (2004), which exploits changes in BMI due to genetic factors as a source of exogenous variation. Specifically, this author proposes the use of the BMI of a sibling as instrumental variable for BMI, arguing that the BMI between siblings is strongly correlated because they share much of the same genes. The crucial assumption of this identification strategy is that the BMI of a sibling is not correlated with unobserved determinants of labor market outcomes. The plausibility of this assumption relies on the evidence suggesting that the correlation between BMI across siblings is not explained by home environment (COMUZZIE; ALLISON, 1998; VOGLER et al., 1995).

Differently from previous studies, we extend the analysis to a measure related to employment quality. Specifically, we investigate the effect of body size on labor informality among

Brazilian workers. The literature has emphasized the potential detrimental effects of informality on well-being due to inferior working conditions. If the BMI is related to productivity, then differences in BMI could explain variations in the employment conditions because the least productive individuals may be unable to find work in the formal sector. Alternatively, less productive individuals may see the informal sector as a mechanism for receiving higher income than that they would receive in the formal sector.

Brazil is a developing country characterized by one of the highest inequalities in income. At the same time, Brazil has a high fraction of people living below the poverty line. The figures indicate that about 6.8 % of the population living on less than \$2.00 a day in 2012(WORLD BANK, 2015). Not surprisingly, in many other ways Brazil is a developing country: life expectancy is 73 years, which is nearly 15 years lower than in countries such as Japan; infant mortality is 12 per hundred thousand inhabitants, a figure 150% higher than the corresponding to US; and the incidence of malnutrition is 10 times higher than in countries like Canada. Therefore, an investigation of the magnitude of the effect of body weight on labor market outcomes would provide useful information for the formulation of policies aimed at reducing inequality and poverty in the country.

To preview, we found that increases in BMI is associated with better labor market outcomes. Specifically, we found that BMI has a positive effect on wages and reduces the likelihood of informal employment (measured as lack of social security). We also find heterogeneous effects by sex, race and urban/rural status. Indeed, we find a greater impact on women, non-white people and individuals living in urban areas. We also found suggestive evidence that education is an important mechanism through which BMI affects labor market outcomes. In particular, we find that the effect of BMI on wages is reduced by about 80% when the years of schooling is included as a control variable. Education is a plausible mechanism in view of studies showing that poor health, particularly during early stages of life, may adversely affect human capital accumulation(CURRIE, 2009; JANET CURRIE; ENRICO MORETTI, 2007). Thus, one possible interpretation is that lower BMI means poorer health, which implies less human capital formation and, in turn, worse labor market outcomes. However, we recognize the possibility that our analysis underestimate or exaggerate the role of education as mechanisms since we do not solve the endogeneity of education. Therefore, we see this as an exploratory exercise and should be interpreted as suggestive.

The rest of the paper is organized as follows. The following section presents a review of the literature. The third section presents a description of the data. The fourth section presents a description of the methodology used and discusses the plausibility of using the strategy proposed by Cawley (2004) in the Brazilian context. In the fifth section, we present the results of the estimates from OLS and Instrumental Variables (IV). Finally, the sixth section concludes with a discussion of the results.

2. Literature

There is a large body of studies investigating how labor market outcomes respond to variations in body weight. Most of this empirical evidence is concentrated in developed countries⁵. In particular, the relationship between weight and wages tend to be negative in developed world, but positive in developing countries(CAWLEY; HAN; NORTON, 2009). For developing countries, it has been documented that the increase of weight (from underweight to normal) is associated with the improvement of current health status and thus with higher level of labor productivity (SCHULTZ; PETERS; HEDLEY, 1995; THOMAS; STRAUSS, 1997). Another possible explanation for the negative sign in developing economies is attributed to the fact that jobs are

⁵ See, for example, Sargent and Blanchflower (1994), Baum and Ford (2004), Shimokawa (2008), Greve (2008), Kline and Tobias (2008) and Norton and Han (2008).

more exhausting in developing countries and higher body mass upgrades productivity in this way (GLICK; SAHN, 1998).

In an often-cited study, Cawley (2004) found that weight has negative impact on wages for white females using employs the fixed effect and IV models with the BMI of a sibling as instrument for BMI in order to deal with the endogeneity of weight. The authors pointed out that OLS estimates for those females indicate that difference in weight of two standard deviations (around 64 pounds) is associated with a difference in wages of 9 percent. This difference in wages is equivalent in absolute value to the wage effect of roughly 1.5 years of education or three years of work experience (CAWLEY, 2004, p 468). His results of other gender-ethnic groups showed a negative weight–wage relationship caused by unobserved heterogeneity. Other empirical contributions have already indicated that heavier individuals tend on average to earn less in developed countries (KLINE; TOBIAS, 2008; NORTON; HAN, 2008). For example, Conley and Glauber (2006) using data from the Panel Study of Income Dynamics (PSID), estimate sibling fixed effects models where a body mass index measure is lagged by 15 years to correct for endogeneity bias. Conley and Glauber (2006) compare OLS estimates with sibling fixed effects estimates and find that obesity is associated with an 18% reduction in women's wages and a 16% reduction in women's probability of marriage. Using data on bioelectrical impedance analysis for United States, Wada and Tekin (2010) develop measures of body composition, body fat, and fat-free mass, and analyze the relationship with wages. Their results indicate that body fat is associated with decreasing wages for both women and men, and fat free mass is associated with increased wages. Sabia and Rees (2012) used OLS and fixed effects estimates to provide evidence that overweight white women are paid substantially less per hour than their slimmer counterparts. Finally, Sabia and Rees (2012) used a model two-stage least squares (2SLS) that used sibling BMI or mother's obesity status in order to confirm and isolate exogenous variation in body weight.

Conversely to studies described above, literature in developing countries seems to show a positive sign in the correlation between weight and wage⁶. In view, Schultz (2003) found that a unit increase in BMI is associated with a 9% raise in wages for women in the Ivory Coast. His results for men suggested that an additional unit of BMI increases men's wages by 15% in the Ivory Coast and by 7% in Ghana. Schultz (2003) used instrumental variable techniques to correct for the potential endogeneity and errors in measurement of BMI. Dinda et al., (2006), using data from of Indian coalmine workers, estimate the associations between body weight and labor market outcomes. Dinda et al., (2006) found a positive effect of BMI on wage earnings. For Brazil, Thomas and Strauss (1997) examined the link of body side with wages among a sample that gathered detailed information of 53,000 households on incomes, expenditures, and socio-demographic characteristics of household members. Using relative prices of food as instrumental variables for BMI, these authors provided evidence that BMI is associated with higher wages for males after controlling for height and nutrient intakes, and these impacts are generally largest among the less educated.

3. Data

This study uses data from the last Consumer Expenditure Survey – POF (*Pesquisa de Orçamentos Familiares*) 2008-2009 which was conducted by the Brazilian Census Bureau – IBGE (*Instituto Brasileiro de Geografia e Estatística*) between May 2008 and May 2009 through interviews in a sample of Brazilian households. Each household was interviewed for a 7-day period through questionnaires regarding the household and its members. The survey aimed to provide information on the household budget composition and on population's life conditions as to measure the structures of consumption, expenditures and income sources. The sampling of

⁶ For instance, a higher weight-for-height is associated with higher wages in Philippines (HADDAD; BOUIS, 1991) and Guinea (GLICK; SAHN, 1998).

POF is part of the group of complex surveys which involves the use of three types of probability samples (simple random sampling, stratified sampling and cluster sampling). Firstly, census sectors, which are geographical areas defined in the Brazilian 2000 Demographic Census, were selected to form a chief sample used by IBGE in all household surveys of the bureau. These sectors were then divided in strata by a geographical and statistical stratification method⁷. Secondly, census sectors were sampled from each stratum and formed a subsample used to select the 55,970 households of POF final sample. All the information is available in the form of microdata disposed in different records. Only Registers 02 and 14 were used in this study, the first is related to the people living in the household and the second has information on labor, income and income deductions. The sample of this study includes family⁸ members in the household from 18 to 40 years old and was restricted to individuals that were registered as children in POF. In other words, these individuals are related to the household manager through a parental relationship. This restriction was imposed due to the estimation method chosen to test the hypothesis on how Body Mass Index (BMI) affects household members' labor market outcomes. More precisely, the estimation requires information on the individual's siblings. Furthermore, the sample excluded individuals that were not in the labor market.

In Brazil, country-representative household budget surveys have been conducted since the seventies' decade and more recently information on anthropometric measures was collected from each member in the household during the period of the interview. These measures contain information such as body weight and height which made possible to calculate the BMI. This index was then estimated for each individual by dividing weight (kilos) by the square of height (meters).

In order to establish a relationship between BMI and labor market outcomes the variables considered were people's hourly wages and a proxy created to measure informality in the labor market. The hourly wage variable is simply the division of the individual's monthly labor income by the number of monthly hours worked. We only considered in the analysis the main job declared by the respondent and jobs of only four groups: private sector, public sector, domestic and temporary (rural areas). In addition, we only considered monthly wages above 100.00 Brazilian *Reais* to generate the hourly wage variable. To create the proxy for informality, we used one of the indexes presented Henley, Arabsheibani, and Carneiro (2009). These authors define informality as the number of people in the labor market that do not contribute to any form of social security. We created a binary variable as a measure of the informal labor market; this variable is equal to 1 if the individual did not contribute to a public social security system, and 0, otherwise. The other variables included were the individual's social and demographic characteristics such as age, month of birth, sex, race, education years, area of residence and region. These are considered control variables that could also affect labor market outcomes.

Table 1 presents the summary statistics for all the variables discussed. It is interesting to notice the high participation of informal labor market at 42%. The maximum hourly wage observed was 116.79 Brazilian *Reais*. The level of BMI is on average 23.67 which is considered normal weight according to World Health Organization classification⁹ (WHO, 1995). The age is on average 26 years and the years of education is on average 10 years. About 58% of the individuals are male and 52% declared themselves as whites. Only 9% of the individuals live in rural areas while 41% belong to cities in the metropolitan area. Most of the individuals live in the Southeast of Brazil (53%) followed by Northeast (21%) and South (14%).

⁷For more details on POF stratification method see the publication "*Despesas, rendimentos e condições de vida*" of IBGE.

⁸ The definition of family differs from the definition of household in POF. One household is composed by people who live under the same roof and share food and living expenses while families are blood-related individuals or individuals connected by domestic dependence or living rules. This implies that one household can have more than one family. Particularly, this condition was observed in only 5.34% of the 55,970 sampled households.

⁹ The WHO classification of body weight is: underweight if $BMI < 18.5 \text{ kg/m}^2$, normal if $18.5 \text{ kg/m}^2 \leq BMI < 25 \text{ kg/m}^2$, overweight if $25 \text{ kg/m}^2 \leq BMI < 30 \text{ kg/m}^2$ and obese if $BMI \geq 30 \text{ kg/m}^2$.

4. Empirical Strategy

We use the technique of instrumental variables to estimate the effect of BMI on labor market outcomes as follows:

First stage:

$$BMI_i = \theta BMI_i^{sibling} + X'\beta + \xi_i \quad (1)$$

Second stage:

$$y_i = \delta \widehat{BMI}_i + X'\pi + \eta_i \quad (2)$$

BMI is the body mass index of the individual i and $BMI_i^{sibling}$ is the body mass index sibling. y represents different measures of labor market outcomes. The X vector includes a set of specific individual characteristics such as sex, race, and civil status. The terms ξ and η are idiosyncratic errors. The standard errors are clustered at States level.

The final coefficient of interest is δ , which represents the effect of BMI on a given labor market outcome measure. The model essentially follows the same strategy of identifying of Cawley(2004). The assumptions behind the identification strategy are: i) the sibling BMI has a significant correlation with the BMI; and ii) the sibling BMI is properly excluded from the equation (2), or equivalently is uncorrelated with η . The validity of the first assumption is expected due to the sibling of same parents share half of their genes. This would ensure a strong correlation between BMI and sibling BMI.

The second assumption is more challenging since we have no way to test their validity in the data. However, there is indirect evidence for this assumption, as discussed by Cawley (2004). However, it could be argued that non-genetic variation in BMI could be correlated with the error term of the second-stage. In particular, a possibility is that learning habits at home are correlated with BMI and wages. Nevertheless, there is no evidence finding that a significant effect of common environment within the home on BMI (COMUZZIE; ALLISON, 1998; VOGLER et al., 1995). On the other hand, there is evidence shown, for example, that the BMI between adopted siblings are significantly correlated. This suggests that most of the BMI correlation between siblings is mainly explained by genetic factors. These facts make plausible the main assumption of the identification strategy.

5. Results

5.1. OLS and IV Estimates

The first procedure to analyze the relationship between BMI and labor market outcomes was to estimate OLS regressions (See Table 2). Panel A shows the results for wage while panel B shows the results for informal employment. Column (1) is based on a specification that includes State and Cohort fixed effects (dummies controlling for the individual's birth year) and a dummy for male individuals as control variables. The other columns correspond to different specifications that include additional groups of controls. At the bottom of the table, the number of observations presented.

The BMI parameter varies from a significant effect on the logarithm of hourly wages in the first specification to insignificant estimates in the remaining ones. In terms of magnitude, the parameter of interest suffered a substantial decrease when more control variables were included. When the controls related to the area of residence are added the coefficient of interest is reduced by 32%, and now is not significant. Indeed, when all the control variables are added (column 4), the coefficient of interest is reduced by 223%.

On the other hand, the relationship between BMI and informal employment is never statistically significant. As more groups of controls are included, the parameter became larger. It is also important to note that the sign of the coefficients is positive; it suggests that higher BMI increases the likelihood of informal employment. This is counterintuitive.

In general, there is little evidence of a systematic association between BMI and labor market outcomes. In addition, the results are very sensitive to control variables included in the regressions. This suggests that the measurement error in BMI and omitted variable bias could play an important role.

The second approach to estimate the relationship between BMI and labor market outcomes in Brazil was carried out by making use of instrumental variables. All the estimates are from two-stage least squares estimation. We opted to use as instrument the BMI of the individual's sibling. Table 3 presents evidence that the sibling's BMI is strongly correlated with BMI. Specifically, the estimates shown are for the equation that relates the individual's BMI to his/her sibling's BMI. The sign of the coefficient is positive and statistically significant in all regressions, as expected. Also, the F-statistics for instrument exclusion above 10 indicates that the sibling's BMI has sufficient predictive power on BMI. Importantly, the parameter of interest does not change significantly, as different controls were included in the regression. For example, the coefficient of interest changes from 0.17 to 0.168. This can be interpreted as evidence favorable to the fundamental assumption of the identification strategy.

Table 4 presents instrumental variables estimates of the association between BMI and labor market outcomes. In this regression the individual's own BMI was instrumented by his/her sibling's BMI and the same controls included in the previous regressions were used as shown in Columns from (1) to (4). In contrast to OLS estimates, the instrumental variables results are all significant. The addition of more sets of controls caused a decrease in the parameters magnitude in both the hourly wage and the informal employment equations. However, it seems to stay more stable in comparison to OLS estimates. The results indicate that increases in BMI are associated with increases in wages. An important difference with the OLS results is that now the effect on informal employment BMI is negative and statistically significant, suggesting that increases in BMI reduces the likelihood of informal employment.

The BMI effect on labor market outcome was also evaluated in different subsamples to capture the heterogeneity of this effect among different groups of the society. Table 5 illustrates how BMI may affect hourly wages and informal employment according to the individuals' sex, race and residence area. It was used instrumental variables estimation considering all the group of controls. Interestingly, the positive impact on wages and the negative impact on informal employment were also observed in all subsamples. Nevertheless, BMI may not affect labor market outcomes of some groups separately.

The parameters were significant only for female individuals, non-white individuals and for people living in urban areas in the hourly wages equation. On the informal employment equation, the significant parameters were also for female individuals and in contrast to the hourly wage equation, informal employment seems to be affected by BMI only for people in rural areas. These results indicate that BMI is positively related to women's hourly wages and it has no influence on male individuals' labor market outcome.

The significant relationship between BMI effect and hourly wages of non-white individuals indicates that race is a relevant aspect on this subject as a higher body weight or a better sign of nutritional condition may enhance higher wages in this specific group. It makes sense that only individuals in urban areas may present a positive response on hourly wages by an increase on BMI. Labor markets in urban areas tend to be more competitive and individuals with a higher body mass may more productive and therefore better compensated.

5.2. *Does Education Matter?*

Our final effort is to present evidence on a potential channel through which BMI may affect labor market outcomes. In particular, we investigate if education plays a role. In order to address

this issue directly, Table 6 present an exercise in where we run the benchmark specification controlling for years of education (Column 2). The education could be a mechanism give that human capital accumulation is directly affected by individual's health conditions, as suggested by the literature (CURRIE, 2009; JANET CURRIE; ENRICO MORETTI, 2007). For comparison proposal, the column (1) of Table 6 presents the results of the column (4) of Table 4. The results in panel A in the table show that the effect of BMI on wages are reduced and no longer statistically significant when education years is included as a control variable. In fact, the effect of BMI on hourly wages is reduced by about 80%. Panel B results indicate that the effect of BMI on informal employment is also reduced (in absolute value) and no longer statistically significant. These results suggest that education may be an important mechanism through which the BMI affects labor market outcomes. However, we emphasize that this exercise can be considered as suggestive and demands further investigation.

5.3. *Robustness of Findings*

Our strategy to test the validity of the results was to analyze the changes on BMI parameters by performing two different specifications. One consisted in the exclusion of outliers in BMI and the other in the inclusion of extreme low values in wages. Table 7 presents the results of the alternative specifications in comparison to the baseline model in Column 1. The results in the table show that the estimates are not sensible to outliers both in BMI and wage.

6. Final Remarks

Analyzing the relationship between Body weight and labor market outcomes has been the main subject of this study. To deal with the potential endogeneity of body weight, we use instrumental variable technique. Our findings suggest that higher BMI is associated with higher wages and reductions in the probability of informal employment. In addition, we found that the effects are greater in women, the nonwhite people and individuals living in urban areas. These results contrast with those of developed countries where a negative relationship is found between BMI and labor market outcomes. One potential explanation is that increments in BMI mean improvements in the health and nutrition due to that of nutrient intake is lower in developing countries. This interpretation is also consistent with studies that show a nonlinear association between BMI and wages. Specifically, these studies show that BMI follow an inverse-U-shaped pattern relative to wages, being the turning point a BMI of 29 (SHIMOKAWA, 2008; WADA; TEKIN, 2010).

Our results also suggest that education may be an important mechanism through which the BMI affects the results of the labor market. This is plausible in view of studies showing that early nutrition is an important determinant for formation of human capital. Therefore, our results could represent the long-term relationship between BMI and labor market outcomes. Nonetheless, evidence of this study is essentially suggestive and requests further research.

We are aware that there are several limitations. The use sibling's BMI as an instrument implies that our sample is based on individuals who have at least one sibling. Therefore, our results may not be generalized to the entire population. Furthermore, the use of that instrument necessarily implies a reduction in the number of observations which could affect the precision of estimates. Setting aside the issue of generalizability, the fact that the performance in the labor market is directly related to poverty and individual welfare, estimating the causal effect of malleable variables in terms of policy as the BMI is an important question *per se*.

References

ABBOTT, R. D. et al. Body mass index and thromboembolic stroke in nonsmoking men in older middle age. The Honolulu Heart Program. **Stroke**, v. 25, n. 12, p. 2370–2376, dez. 1994.

BAUM, C. L.; FORD, W. F. The wage effects of obesity: a longitudinal study. **Health Economics**, v. 13, n. 9, p. 885–899, 2004.

CAWLEY, J. The Impact of Obesity on Wages. **The Journal of Human Resources**, v. 39, n. 2, p. 451–474, 1 abr. 2004.

CAWLEY, J.; HAN, E.; NORTON, E. C. Obesity and labor market outcomes among legal immigrants to the United States from developing countries. **Economics & Human Biology**, v. 7, n. 2, p. 153–164, jul. 2009.

COMUZZIE, A. G.; ALLISON, D. B. The search for human obesity genes. **Science**, v. 280, n. 5368, p. 1374–1377, maio 1998.

CONLEY, D.; GLAUBER, R. Gender, Body Mass, and Socioeconomic Status: New Evidence from the PSID. In: **The Economics of Obesity**. Advances in Health Economics and Health Services Research. [s.l.] Emerald Group Publishing Limited, 2006. v. 17p. 253–275.

CURRIE, J. Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. **Journal of Economic Literature**, v. 47, n. 1, p. 87–122, mar. 2009.

CURRIE, J.; VOGL, T. Early-Life Health and Adult Circumstance in Developing Countries. **Annual Review of Economics**, v. 5, n. 1, p. 1–36, jan. 2013.

GLICK, P.; SAHN, D. E. Health and productivity in a heterogeneous urban labour market. **Applied Economics**, v. 30, n. 2, p. 203–216, 1 fev. 1998.

HADDAD, L. J.; BOUIS, H. E. The impact of nutritional status on agricultural productivity: wage evidence from the Philippines. **Oxford bulletin of economics and statistics**, v. 53, n. 1, p. 45–68, 1991.

HENLEY, A.; ARABSHEIBANI, G. R.; CARNEIRO, F. G. On Defining and Measuring the Informal Sector: Evidence from Brazil. **World Development**, v. 37, n. 5, p. 992–1003, maio 2009.

JANET CURRIE; ENRICO MORETTI. Biology as Destiny? Short and Long Run Determinants of Intergenerational Transmission of Birth Weight. **Journal of Labor Economics**, v. 25, n. 2, p. 231–264, abr. 2007.

JD, S.; DG, B. Obesity and stature in adolescence and earnings in young adulthood: Analysis of a british birth cohort. **Archives of Pediatrics & Adolescent Medicine**, v. 148, n. 7, p. 681–687, 1 jul. 1994.

KLINE, B.; TOBIAS, J. L. The wages of BMI: Bayesian analysis of a skewed treatment–response model with nonparametric endogeneity. **Journal of Applied Econometrics**, v. 23, n. 6, p. 767–793, 1 set. 2008.

NORTON, E. C.; HAN, E. Genetic Information, Obesity, and Labor Market Outcomes. **Health economics**, v. 17, n. 9, p. 10.1002/hec.1383, set. 2008.

PI-SUNYER, F. X. The obesity epidemic: pathophysiology and consequences of obesity. **Obesity research**, v. 10, n. 2, p. 97–104, dez. 2002.

SABIA, J. J.; REES, D. I. Body weight and wages: Evidence from Add Health. **Economics & Human Biology**, v. 10, n. 1, p. 14–19, jan. 2012.

SCHULTZ, T. P. Wage rentals for reproducible human capital: evidence from Ghana and the Ivory Coast. **Economics & Human Biology**, v. 1, n. 3, p. 331–366, 2003.

SCHULTZ, T. P.; PETERS, G. H.; HEDLEY, D. D. **Human capital and economic development**. Dartmouth Publishing Co. Ltd., 1995

SHIMOKAWA, S. The labour market impact of body weight in China: a semiparametric analysis. **Applied Economics**, v. 40, n. 8, p. 949–968, 2008.

THOMAS, D.; STRAUSS, J. Health and wages: Evidence on men and women in urban Brazil. **Journal of Econometrics**, v. 77, n. 1, p. 159–185, mar. 1997.

VOGLER, G. P. et al. Influences of genes and shared family environment on adult body mass index assessed in an adoption study by a comprehensive path model. **International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity**, v. 19, n. 1, p. 40–45, jan. 1995.

WADA, R.; TEKIN, E. Body composition and wages. **Economics & Human Biology**, v. 8, n. 2, p. 242–254, jul. 2010.

WORLD HEALTH ORGANIZATION. Expert Committee. Physical Status: The Use and Interpretation of Anthropometry. **WHO technical report** series no 854. Geneva: WHO ed, 1995.

WORLD BANK. World Development Indicators 2015. Washington, DC. © **World Bank**. <https://openknowledge.worldbank.org/handle/10986/21634> License: CC BY 3.0 IGO,2015

Tables

Tables

Table 1. Summary Statistics

	N	Mean	Standard Deviation	Min	Max
<i>Labor Market Outcomes:</i>					
Hourly wage	8,831	4.83	5.60	0.47	116.79
Informal employment	8,461	0.42	0.49	0	1
<i>Socioeconomic and Demographic Characteristics</i>					
Body Mass Index (BMI)	8,831	23.67	4.32	13.80	150
Age (in years)	8,831	26.06	7.57	18	60
Male	8,831	0.58	0.49	0	1
White	8,831	0.52	0.50	0	1
Education years	8,770	10.48	3.14	0	15
<i>Area residence :</i>					
Rural	8,831	0.09	0.28	0	1
Metropolitan City	8,831	0.41	0.49	0	1
<i>Region:</i>					
North	8,831	0.05	0.21	0	1
Northeast	8,831	0.21	0.41	0	1
Southeast	8,831	0.53	0.50	0	1
South	8,831	0.14	0.35	0	1
Midwestern	8,831	0.07	0.25	0	1

Source: Research results.

Notes: All statistics were obtained using sample weights.

Table 2. OLS Estimates of the Relationship between BMI and Labor Market Outcomes

	(1)	(2)	(3)	(4)
<i>Panel A: Log of Hourly Wage</i>				
Body Mass Index	0.00356 [0.00192]*	0.00268 [0.00184]	0.00240 [0.00196]	0.00110 [0.00175]
Number of observations	8,815	8,815	8,815	8,754
<i>Panel B: Informal Employment</i>				
Body Mass Index	0.000635 [0.00179]	0.000880 [0.00174]	0.00112 [0.00175]	0.00149 [0.00183]
Number of observations	8,446	8,446	8,446	8,386
State Fixed Effects	Yes	Yes	Yes	Yes
Cohort Fixed Effects	Yes	Yes	Yes	Yes
Male Dummy	Yes	Yes	Yes	Yes
Residence Area	No	Yes	Yes	Yes
Other Demographic Characteristics	No	No	Yes	Yes
Household Head Characteristics	No	No	No	Yes

Source: Research results.

Notes: Standard errors (in brackets) are robust to heteroskedasticity. State fixed effects include dummies of State of residence. Cohort fixed effects include birth year dummies. Residence area includes dummies indicating metropolitan city and rural residence status. Other demographic characteristics contain dummies indicating birth month and white race. Household head characteristics are education years, age, and sex. All estimations use sample weights. * $p < 0.1$.

Table 3. First Stage Results: Relationship between BMI and Sibling BMI

	<i>Dependent variable is BMI</i>			
	(1)	(2)	(3)	(4)
Sibling Body Mass Index	0.17 [0.0240]***	0.17 [0.0243]***	0.169 [0.0244]***	0.168 [0.0249]***
Exc. instruments F-stat.	50.11	48.84	48.18	45.5
Kleibergen and Paap test (p-value)	0.00	0.00	0.00	0.00
Number of observations	8,781	8,781	8,781	8,725
Basic Controls	Yes	Yes	Yes	Yes
Residence Area	No	Yes	Yes	Yes
Other Demographic Characteristics	No	No	Yes	Yes
Household Head Characteristics	No	No	No	Yes

Source: Research results.

Notes: Standard errors (in brackets) are robust to heteroskedasticity. Basic controls include State and cohort fixed effects, sex and dummies for sibling's birth year and sex. Residence area contains dummies indicating metropolitan city and rural residence status. Other demographic characteristics contain dummies for birth month and white race, and sibling's month birth and white race. Household head characteristics are education years, age, and sex. All estimations use sample weights. *** p<0.01.

Table 4. Instrumental Variables Estimates of the Relationship between BMI and Labor Market Outcomes

	(1)	(2)	(3)	(4)
<i>Panel A: Log of Hourly Wage</i>				
Body Mass Index	0.0342 [0.0141]**	0.0318 [0.0132]**	0.0264 [0.0130]**	0.0248 [0.0116]**
Number of observations	8,781	8,781	8,781	8,725
<i>Panel B: Informal Employment</i>				
Body Mass Index	-0.0180 [0.00943]*	-0.0165 [0.00947]*	-0.0161 [0.00908]*	-0.0161 [0.00912]*
Number of observations	8,413	8,413	8,413	8,359
Basic controls	Yes	Yes	Yes	Yes
Residence Area	No	Yes	Yes	Yes
Other Demographic Characteristics	No	No	Yes	Yes
Household Head Characteristics	No	No	No	Yes

Source: Research results.

Notes: Standard errors (in brackets) are robust to heteroskedasticity. Basic controls include State and cohort fixed effects, sex and dummies for sibling's birth year and sex. Residence area contain dummies indicating metropolitan city and rural residence status. Other demographic characteristics contain dummies for birth month and white race, and sibling's month birth and white race. Household head characteristics are education years, age, and sex. All estimations use sample weights. * p<0.1, ** p<0.05, *** p<0.01.

Table 5. Instrumental Variables Estimates of the Relationship between BMI and Labor Market Outcomes by Sex, Race and Residence Area

	Male (1)	Female (2)	White (3)	Non-White (4)	Urban (5)	Rural (6)
<i>Panel A: Log of Hourly Wage</i>						
Body Mass Index	0.0159 [0.0145]	0.0326 [0.0178]*	0.0182 [0.0202]	0.0381 [0.0179]**	0.0262 [0.0125]**	0.00384 [0.0275]
Number of observations	5,133	3,592	3,779	4,946	7,485	1,240
<i>Panel B: Informal Employment</i>						
Body Mass Index	-0.000759 [0.0120]	-0.0434 [0.0157]***	-0.0168 [0.0137]	-0.0179 [0.0141]	-0.0123 [0.00930]	-0.0459 [0.0212]**
Number of observations	4,901	3,458	3,619	4,740	7,151	1,208
Basic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Residence Area	Yes	Yes	Yes	Yes	Yes	Yes
Other Demographic Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Household Head Characteristics	Yes	Yes	Yes	Yes	Yes	Yes

Source: Research results.

Notes: Standard errors (in brackets) are robust to heteroskedasticity. Basic controls include State and cohort fixed effects, sex and dummies for sibling's birth year and sex (columns 1 and 2 do not include sex, while columns 3 and 4 do not include white dummy). Residence area contain dummies indicating metropolitan city and rural residence status (columns 5 and 6 do not include rural dummy). Other demographic characteristics contain dummies for birth month and white race, and sibling's month birth and white race. Household head characteristics are education years, age, and sex. All estimations use sample weights. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6. Instrumental Variables Estimates of the Relationship between BMI and Labor Market Outcomes (Controlling by Education Years)

	Baseline (1)	Controlling by Education years (2)
<i>Panel A: Log of Hourly Wage</i>		
Body Mass Index	0.0248 [0.0116]**	0.0137 [0.0106]
Education Years		0.0798 [0.00361]***
Number of observations	8725	8668
<i>Panel B: Informal Employment</i>		
Body Mass Index	-0.0161 [0.00912]*	-0.0121 [0.00877]
Education Years		-0.0318 [0.00286]***
Number of observations	8359	8302
Basic Controls	Yes	Yes
Residence Area	Yes	Yes
Other Demographic Characteristics	Yes	Yes
Household Head Characteristics	Yes	Yes

Source: research results.

Notes: standard errors (in brackets) are robust to heteroskedasticity. Basic controls include State and cohort fixed effects, sex and dummies for sibling's birth year and sex. Residence area contain dummies indicating metropolitan city and rural residence status. Other demographic characteristics contain dummies for birth month and white race, and sibling's month birth and white race. Household head characteristics are education years, age, and sex. All estimations use sample weights. * p<0.1, ** p<0.05, *** p<0.01.

Table 7. Instrumental Variables Estimates of the Relationship between BMI and Labor Market Outcomes (Robustness of Findings)

	Baseline	Excluding Outliers in BMI	Including extreme low wages
	(1)	(2)	(3)
<i>Panel A: Log of Hourly Wage</i>			
Body Mass Index	0.0248 [0.0116]**	0.0359 [0.0151]**	0.0298 [0.0119]**
Number of observations	8,725	8,565	9,072
<i>Panel B: Informal Employment</i>			
Body Mass Index	-0.0161 [0.00912]*	-0.0219 [0.0118]*	-0.0177 [0.00915]*
Number of observations	8,359	8,208	8,727
Basic controls	Yes	Yes	Yes
Residence Area	Yes	Yes	Yes
Other Demographic Characteristics	Yes	Yes	Yes
Household Head Characteristics	Yes	Yes	Yes

Source: Research results.

Notes: Standard errors (in brackets) are robust to heteroskedasticity. Column 2 eliminates potential outliers by restricting our sample to include only individuals with a BMI above 15 and lower than 35. Column 3 includes individuals with a wage lower than 100.00BrazilianReais in our sample. Basic controls include State and cohort fixed effects, sex and dummies for sibling's birth year and sex. Residence area contain dummies indicating metropolitan city and rural residence status. Other demographic characteristics contain dummies for birth month and white race, and sibling's month birth and white race. Household head characteristics are education years, age, and sex. All estimations use sample weights. * p<0.1, ** p<0.05, *** p<0.01.

Secondary vocational education and earnings in Brazil

Marina Aguas (IBGE, Rio de Janeiro, Brazil, e-mail: mariffaguas@gmail.com)
Danielle Carusi Machado (UFF, Niterói, Brazil, e-mail: daniellecariusi@vm.uff.br)

Abstract

Vocational education is a type of human capital investment that aims to provide workers with knowledge, know-how, skills, specific competences and practices in a specific area, which is expected to improve their performance in the labor market. Investing in vocational education can also reduce the barriers to an individual's entry in their first-time job, as well as improve the quality of the job-worker match. Consequently, more skilled and productive workers can benefit from higher wages, promoting positive impacts for the entire economy. This paper investigates whether secondary vocational education is related to labor earnings in Brazil. Making use of data from the 2007 PNAD, the empirical analysis is conducted through three different methodologies: OLS, Treatment Effect and Propensity Score Matching. Evidence shows that attainment of a secondary vocational education is associated with an increase in labor earnings between 20% and 24% in all these methods. Evidence also suggests that the decision to attend such programs does not seem to be correlated with unobserved productive characteristics of workers.

Keywords: labor market, income, vocational training.

JEL: C21, J24, J31

Resumo

Educação profissional é um tipo de investimento em capital humano que tem como objetivo prover os trabalhadores com conhecimentos e competências/práticas específicas para determinada área, sendo esperado que isto melhore sua performance no mercado de trabalho. Investir em educação profissional pode igualmente reduzir as barreiras de entrada ao primeiro emprego e, também melhorar o casamento entre trabalhadores/empregadores. Consequentemente, trabalhadores mais produtivos e com melhores conhecimentos ganham salários mais altos, impactando positivamente a economia. O objetivo deste artigo é investigar como a educação profissional se relaciona com os rendimentos do trabalho no Brasil. Usaremos os dados da PNAD de 2007 e adotaremos três metodologias empíricas: OLS, efeito tratamento e PSM. Mostramos que o impacto da educação profissional é de aumentar os ganhos de 20% a 24% em todos métodos e que a decisão de realizar estes cursos não parece ser correlacionada com características produtivas não observadas.

Palavras chaves: mercado de trabalho, educação profissional, renda do trabalho.

Área 2: Economia social e do trabalho

1. Introduction

Vocational education is a type of human capital investment that aims to provide workers with knowledge, know-how, skills, specific competences and practices in a specific area, which is expected to improve their performance in the labor market. Investing in vocational education can also reduce the barriers to an individual's entry in their first-time job, as well as improve the quality of the job-worker match. Consequently, more skilled and productive workers can benefit from higher wages, promoting positive impacts for the entire economy.

In Brazil, the current education law (# 9394/1996) defines three types of vocational education: vocational training, secondary vocational education and tertiary vocational education. Each one of these categories is structured for a specific target, with specific prerequisites and objectives. These programs usually have a shorter duration than regular education courses. Vocational education can be seen as a faster way of qualifying the workforce, especially in countries with a large population characterized by low levels of education.

The aim of this paper is to investigate the relationship between earnings and vocational education. Specifically, we chose the secondary vocational education (SVE) to do our analysis, because it is considered an intermediate level of the vocational education, it has guidelines and rules for all courses defined by Brazilian law and it is not a tertiary education course. Therefore, in order to receive the secondary vocational education degree, students must have finished the secondary school, as well as attend and successfully complete vocational courses recognized and structured by the Brazilian government. Thus, this segment differs from both vocational training courses, which are not regulated and have no educational prerequisites, and tertiary vocational education, which is part of the undergraduate courses in Brazil.

The empirical analysis is inspired by Mincer (1974), which defines an earnings equation as a function of educational level, experience and individual characteristics. After some adjustments, our analysis incorporates the SVE as an explanatory variable in the educational context. We use the data from the Brazilian National Household Survey (PNAD) for the year of 2007, as this year was held an additional questionnaire about vocational education. We

chose a sub-sample of people aged between 25 and 55 years old, prime age group for labor market analysis, that live in metropolitan regions or self-representative areas in Brazil.

The estimation of the earnings equation is performed through three different methods: Ordinary Least Squares (OLS), Treatment-Effect Model (TE) and Propensity Score Matching (PSM). These last two methods are proposed to fix the possible selection bias present in the fact that educational choices are not random. This correction is based on observable factors, to the extent that these two methods cannot capture the observed selectivity which is the classical problem of the relationship between education and earnings. The OLS is the usual method and the results are compared with other two methods.

The results show that people who attended a SVE course have a positive and significant impact on their salary per hour, regardless of the estimation techniques used. For the OLS estimation, this effect increases wage per hour in 22%. In the TE method, the expected gain in hourly wage between those who have the training is 21% higher than those who did not. In this method, the results suggest the absence of sample selection bias. Finally, the PSM method shows salary increases ranging between 22% and 24%, according to the matching method chosen. These values are similar to those found in the OLS method and support the idea of the absence of selectivity in the observable characteristics of the sample.

2. Data and descriptive statistics

From the PNAD, of 2007, we selected a sample of employed people, with positive earnings, aged between 25 and 55 years old and with a high school diploma, only. This group should live in metropolitan regions or self-representative areas in Brazil. These choices were made to reduce heterogeneity problems between individuals and regions. The restriction to analyzing people who completed high school but did not attend college, aims to evaluate the return of the training between people who might benefit from it and that have similar characteristics. This sub-sample corresponds to 45% of the population between 25 and 55 years old, in 2007.

With the vocational education questionnaire in PNAD of 2007, it was possible to identify people who attended or completed a SVE course as well as their individual characteristics (sex, age, race, place of residence). Additionally, as part of the TE model, we created two identification variables that are based on observable characteristics of individuals in the sample to try to correct the selectivity bias of the data. These variables attempt to capture if

the place of residence, and its surroundings, influence people's decision to have a training course. Table 1 presents a description of these variables.

Table 1: Variables

Variables	Description
Attended SVE course	Equal to 1 if the person held a SVE course, 0 otherwise.
Complete SVE course	Equal to 1 if the person complete a SVE course, 0 otherwise.
Percentage of people who have completed a SVE course in the stratum of residence	For every person i , aged J and resident in stratum T , calculate the number of people, who are residents in stratum T , aged equal to or greater than J , excluding itself, which concluded a SVE course, divided by the number of people, aged equal to or greater than J , excluding i itself, resident in stratum T .
Percentage of people who did not attend a SVE course because there was no provision of such courses in the stratum of residence	For every person i , aged J and resident in stratum T , calculate the number of people, who are residents in stratum T , aged equal to or greater than J , excluding itself, which did not attend a SVE course because there was no provision of such courses in the stratum of residence, divided by the number of people, aged equal to or greater than J , excluding i itself, resident in stratum T .

With respect to characteristics of the sample, Tables 2 and 3 show the individuals' profiles. Note that 14% of people aged 25-55 years old and with completed secondary school had attended a SVE course, and 12.7% completed this course. Among people who completed SVE course, 35% have never worked in the professional field of training. With respect to individuals without SVE course, it is important to emphasize that most part had no interest in doing them and a minority failed to do so due to the absence of courses.

The individual characteristics of people with and without SVE courses, in Table 3, reveal themselves very close: the percentage of men and women is about the same, the percentage of whites and the average age is slightly larger for those who attended SVE courses. With respect to labor market indicators, we notice some differences. The proportion of employed people with positive labor earnings is 10 percentage points higher for those who attended SVE courses compared to those without such training. The average hourly income for people who did not attend a SVE course is approximately 70% of the average hourly income of those who attended.

Table 2: individual characteristics

Variável	Brasil
People between 25-55 years old who are employed and with positive hourly earnings (people in 1000)	31,417
Attended SVE course (% of people)	13.9
Completed the course	91.4
Completed the course and work or have worked in the training area	64.0
Reason for not attended SVE course: (% of people):	
Is attending now any vocational education course or did a different segment	34.9
no provision of such courses	2.4
Lack of interest	36.8
Other reasons	11.9

Source: own elaboration, using PNAD 2007/IBGE.

Table 3: Characteristic of employed people with positive hourly income by attendance a SVE course

Variables	Attended a SVE course	Did not attend a SVE course
Women	47.3	45.8
White people	57.3	53.5
Average age	37.5	36.4
% of people between 25-55 years old who are employed with positive hourly earnings in the total population of this age group	80.5	70.9
Average hourly income (in BRL from 2007)	8.82	6.21

Source: own elaboration, using PNAD 2007/IBGE.

3. Methodology

Research on the impact of the SVE on labor earnings follows the approach implemented by Mincer (1974) in which labor earnings is represented as a function of education and experience. We try to capture the related effect of carrying out a SVE course on wages including a dummy variable indicating whether the individual attended or not this kind of education in the Mincer equation. Therefore, the model to be estimated is:

$$y_i = \ln R_i = f(nt_i, x_i, z_i) + u_i, i=1, 2, \dots, n \quad (\text{eq.1})$$

Where $\ln R_i$ indicates the natural logarithm of hourly labor earnings of individual i , nt_i is the binary variable indicating whether the person i did or did not a SVE, x_i is her labor experience, z_i are other individual characteristics, like gender, place of residence, etc, and u_i is a stochastic error.

The model is estimated by three methods: ordinary least squares (OLS), treatment-effect (TE) and propensity score matching (PSM). A concern about OLS estimates is the fact that individuals can choose to do or not to do a SVE course and this choice can be a source of bias if some unobservable individual characteristics, such as skills talent, affect wages and also the probability of doing a SVE course. Thus, an individual with more skill talent may receive a higher salary and may also have a higher probability to engage in a vocational course. Therefore, it is possible that higher earnings for those who are trained are due in fact to their greater ability.

In order to dealing with the sample selection problem mentioned before, we use two others methods: TE and PSM, which are very useful in the economic literature of impact evaluation of public policies and social programs. These techniques allow researchers to measure the causal effect of a "treatment" or "program" on a generic outcome variable. In our case, the SVE can be considered a "program" or "treatment" and the amount of hourly labor earnings as the outcome variable. The better way to evaluate the impact of SVE would be to calculate the difference between the individual's labor earnings that result from participating and not in training. We do not observe, however, the wages of individuals who receive training if they had not received it, and we do not observe wages for individuals who were not trained if they were trained. In order to estimate the casual effect of the treatment we should reconstruct the outcomes that are not observable, making use of counterfactuals. The TE and PSM methods propose different ways to estimate counterfactuals outcomes.

3.1. Treatment-effect model¹

This method estimates the effect of nt_i on the continuous variable observed y_i , hourly labor earnings, conditional on exogenous independent variables x_i and w_i , taking into account the decision process of doing or not doing a SVE course. This process is modeled as a result of a latent unobserved variable, nt_i^* , which is supposed to be a linear function of the exogenous explanatory variables w_i , and a random error component e_i . That is,

$$y_i = x_i \beta + \delta nt_i + u_i \text{ (eq. 2)} \quad \quad nt_i^* = w_i \gamma + e_i \text{ (eq. 3)}$$

$$nt_i = \begin{cases} 1, & \text{if } nt_i^* > 0 \\ 0, & \text{otherwise} \end{cases} \quad \text{(eq. 4)}$$

Where, e and u have a bivariate normal distribution with zero mean and correlation equal ρ .

Although this method has a parametric specification and, therefore, the identification can be achieved even if $W = X$, in practice, the estimator will not perform well in the absence of an exclusion restriction in W . The set of variables in X are: sex, age, age squared, race, place of residence and household situation for each individual. The set of variables in W includes those in X and the identification variables shown in the data section.

Identification variables are intended to represent the perceived benefit associated with attending or not a SVE course in the vicinity of the residence of each individual i . We assume that the higher the proportion of persons aged greater than or equal to i who have completed a SVE course, the greater their influence on younger workers decision regarding attending this modality of education. Similarly, higher the percentage of people who did not attained such courses because there was no provision of this type of education in their neighborhood diminishes the younger individuals incentives to attend a SVE course.

Considering the sample selection problem, the wage equation for treated individuals is represented by (5):

$$E(y_{1i} | nt_i = 1) = x_i \beta + \delta + E(u_i | nt_i = 1) = x_i \beta + \delta + \rho \sigma_u \lambda(w_i \gamma) \quad \text{(eq.5)}$$

And the wage equation for non treated individuals is represented by eq. (6):

¹ For more information, see Maddala (1983), Barnowet al. (1981) and Cong&Drukker (2000).

$$E(y_{0i}|nt_i = 0) = x_i \beta + \rho \sigma_u \left[\frac{-\phi(w_i \gamma)}{1 - \Phi(w_i \gamma)} \right] \quad (\text{eq.6})$$

Where λ is the inverse Mills ratio and the parameters will be estimated in two stages.

The difference in expected hourly labor earnings between participants and non-participants can be written as:

$$E(y_{1i}|nt_i = 1) - E(y_{0i}|nt_i = 0) = \delta + \rho \sigma_u \left[\frac{\phi(w_i \gamma)}{\Phi(w_i \gamma)\{1 - \Phi(w_i \gamma)\}} \right] \quad (\text{eq.7})$$

Where ϕ and Φ are density and cumulative distribution functions of a normal distribution. If the correlation between error terms, ρ , is zero, the difference reduces to δ , and we could estimate it using OLS. However, it is believed that ρ is positive, since there are unobserved factors that influence the decision to do a SVE course. Thus, the OLS method should overestimate the treatment effect, δ .

3.2 Matching Model

The Matching model is a non-parametric approach that seeks to reproduce an experiment ex post. The method consists in selecting, among those in the untreated group, a sample of individuals as similar as possible to the treated group, called control group. This method departs from the assumption that all relevant differences between treated and untreated persons are captured by the observed characteristics. Under this assumption, all relevant outcome differences between treated and non-treated individuals are captured by their observable attributes; the only remaining difference between the two groups is the treatment status. In our case, the treatment is the attainment of a SVE course.

The matching on the propensity score is implemented as follows. The method proposes to summarize in a single variable index, called propensity score, the individual characteristics pre-treatment. The propensity score (PS), defined by Rosenbaum and Rubin (1983), is the probability of receiving treatment given the set of observed characteristics jointly affecting treatment status and outcomes: $p(X) \equiv Pr(nt = 1|X) = E(nt|X)$ (eq.8)

Since treated and non-treated individuals with the same value of the PS have the same distribution of the vector of regressors X , and adopting the assumption that treatment is randomly assigned within the groups defined by X , Rosenbaum e Rubin (1983) show that treatment is also randomized within groups defined by $p(X)$. As a result, if the PS of an individual i is known, $p(X_i)$, so the average treatment effect on the treated (ATT) is given by:

$$\begin{aligned}
ATT &\equiv E[Y_{1i} - Y_{0i} | nt_i = 1] = E\{E[Y_{1i} - Y_{0i} | nt_i = 1, p(X_i)]\} \\
&= E\{E[Y_{1i} | nt_i = 1, p(X_i)] - E[Y_{0i} | nt_i = 0, p(X_i)] | nt_i = 1\} \quad (\text{eq.9})
\end{aligned}$$

Where Y_{1i} and Y_{0i} are the potential results of the two contrafactuals situations of treated and non-treated. So, conditioning on the propensity score allows us to obtain unbiased estimates of average treatment effects.

The first step to evaluate SVE effect on earnings is to estimate $\Pr(nt_i=1|X_i)$. We estimate the propensity score using a probit regression model, in which treatment status is regressed on observed baseline characteristics. The estimated PS is the predicted probability of treatment derived from the fitted regression model. As it is too difficult to observe two individuals with the same value of $p(X)$ to do the matching, the match for each individual in treated group to someone in the control group depends on the distance between their propensity scores. We compare four different methods to do this matching: (i) the nearest neighbor matching, which selects the non-participant for whom the value of P_j is closest to P_i , (ii) the radius matching, where a match for individual i is selected only if the propensity score falls within a radius r from P_i , (iii) the kernel matching, where the outcome of treated unit i is matched to a weighted average of the outcomes of more (possibly all) non-treated units, where the weight given to non-treated unit j is a function of the closeness of the characteristics of i and j , and (iv) the stratification method, where the sample is stratified into K groups according to PS intervals. For each group, these covariates are balanced, that is, for each covariate there is no significant difference between the groups.

4. Results

Table 4 shows the OLS estimates results of eq. 1, considering individuals who attended or completed a VSE course. The results show that attending a SVE course means getting an hourly labor earnings about 21% higher than that received by those who have not attended such a course, even with all the prerequisites of doing it. The estimated coefficients are significant at 1% level and vary little when geographical controls are included. For people who finished the course, a slightly higher and significant effect on hourly labor earnings is generated (22%), with our without geographical controls.

Table 4: Results – OLS estimates

Variables	Ln (hourly labor earnings)
-----------	----------------------------

	OLS		OLS with geographical controls	
Attended a SVE course	0.211*** (0.016)		0.213*** (0.017)	
Completed a SVE course		0.221*** (0.017)		0.224*** (0.017)
Observations	22,262	22,262	22,262	22,262
R ²	0.187	0.187	0.212	0.212

Robust standard errors in brackets. *** p<0,01, ** p<0,05, * p<0,1. Source: microdata PNAD 2007 (IBGE).

Table 5 shows EF two-stage estimate results for eq. (2) and (3). For the first stage, coefficients of the identification variables are significant and exhibit the expected sign. The decision to make a SVE course is positively associated with the percentage of people who have completed it in the neighborhood, and negatively related to the percentage of people who did not this type of course because there was no provision of that in the neighborhood. The second stage estimates the wage equation taking into account all the explanatory variables and the selectivity term from the first stage (lambda variable). The coefficients associated with the participation or conclusion of a SVE course are significant and positive. But the selectivity correction term (lambda) is negative and not significant for both specifications. With our two identification variables, sample selection does not seem to exist.

The difference in expected earnings between those who attended the SVE course and those who did not attend it (eq. 7) is significant and equal to 0.207. For those who completed the course, compared to those who did not, the difference is 0.217. Results are slightly lower than those obtained by OLS.

Table 5: Results: TE estimates

Variables	Ln (hourly labor earnings)	Attended a SVE course	hazard	Ln (hourly labor earnings)	Completed a SVE course	hazard
Attended a SVE course	0.343** (0.133)					
Completed a SVE course				0.357** (0.145)		
Percentage of people who have completed a SVE course in the stratum of residence.		0.038*** (0.004)			0.036*** (0.004)	
Percentage of people who did not attend a SVE course because there was no provision of such courses in the stratum of residence		-0.051*** (0.012)			-0.048*** (0.012)	
Lambda			-0.072 (0.073)			-0.073 (0.078)
Constant	0.306*** (0.104)	-1.782*** (0.255)		0.386*** (0.108)	-2.888*** (0.266)	
Observations	22,262	22,262	22,262	22,262	22,262	22,262
Difference between treated and non-treated		0.2074*** (0.010)			0.2176*** (0.010)	

Robust standard errors in brackets. *** p<0,01, ** p<0,05, * p<0,1. Source: microdata PNAD 2007 (IBGE).

Table 6: Propensity Score Matching: estimates

Variables	ATT					
	Attended	# treated	#control	Concluded	# treated	#control
Nearest neighbor matching	0.221 (0.016)	3083	12793	0.231 (0.017)	2793	12443
Kernel	0.232 (0.014)	3083	19151	0.235 (0.015)	2793	19442
Stratification	0.209 (0.001)	3083	19151	0.219 (0.016)	2793	19442
Radiusmatching - radius=0,01	0.229 (0.015)	3083	19100	0.225 (0.015)	2793	19442
Radiusmatching - radius=0,001	0.228 (0.015)	3083	19100	0.220 (0.015)	2792	19372
Radiusmatching - radius=0,0001	0.234 (0.015)	3058	17482	0.229 (0.016)	2775	18020

Source: microdata PNAD 2007 (IBGE).

Table 6 shows the ATT results (eq. 9) using the PSM and the four possible techniques to do the matching between the treatment and the control group. The estimated ATT are positives and significant for all matching methods reported. Estimated coefficients for people who attended the SVE course range between 0.20 and 0.23. Therefore, a person who attended a SVE course has the hourly labor earnings 20% to 23% higher than another one who did not attended the training. The results for people who concluded a SVE course is a little bigger, between 22% to 23,5%.

The four matching methods used to make the correspondence between the treated group and the control group generated similar results. Moreover, results are close to the coefficients obtained by the OLS estimation, which does not consider the presence of sample selection, and also to the results obtained by TE method.

5. Conclusions

Without sample selection, OLS estimates to determine the effects of vocational secondary education on hourly labor earnings should be correct. The OLS results show that attended or concluded a vocational secondary course means getting hourly labor earnings about 21 to 22% higher than that received by those who have not attended such a course. Taking into account the sample selection, TE and PSM methods (and all four techniques to do matching), show similar results to OLS estimates. These estimated differences between treated and non-treated lie in the range 20 to 23%, so, no longer than OLS estimates and suggesting that there is no evidence of sample selection.

Despite this, we must keep in mind that the estimation methods that seek to correct the possible selection bias of the data are based on strong assumptions and their results should be interpreted with caution. In the case of TE method, there is the hypothesis that all heterogeneity among individuals regarding the decision to do a vocational secondary course is captured by the identification variables. In PSM method, conditional on the probability of

participation, the average difference in results between the treatment group and the control is only explained by the participation in a vocational secondary education.

References

Barnow B., Cain G., Goldberger A. ,(1981). Issues in the Analysis of Selection Bias, in the Economics of Training, vol I, Ashenfelter and Lalonde eds., 1996.

Cong, R. and Drukker, D.M. (2000). Treatment Effects Model. [*Stata Technical Bulletin*](#), 2001, vol. 10, issue 55

Maddala, G. S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, New York.

Mincer, J. (1974) Schooling, Experience, and Earnings, National Bureau of Economic Research. Columbia University Press, New York.

Rosenbaum P, Rubin D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983; 70; 41-55.