

O Comunicado do Banco Central*

Rodolfo Cabral[†], Bernardo Guimaraes[‡]

Sumário: 1. Introdução; 2. O comunicado; 3. Dados; 4. Resultados; 5. Conclusão; Apêndice. Valores de comunicados e atas.

Palavras-chave: Banco Central, Comunicado, Comunicação, Política Monetária.

Códigos JEL: E43, E52, E58.

Este artigo estuda como a indicação de política monetária contida nos comunicados do Banco Central do Brasil impacta os mercados. Para isso, classificamos a mensagem de cada nota a partir de uma análise semântica. O comunicado é bastante eficaz enquanto informante do futuro da política monetária, influenciando toda a curva de juros e o Ibovespa e ajudando também a prever o conteúdo da ata de reunião do Comitê de Política Monetária.

This paper studies how the Brazilian Central Bank's statements about its monetary policy decision affect asset prices. The qualitative content of each note is translated into an ordered scale. The statement contains information about future monetary policy decisions: it affects the yield curve and the stock market index, and it helps to predict the content of the minutes of the Monetary Policy Committee meeting.

1. INTRODUÇÃO

O Banco Central do Brasil (BCB) se vale de diversos meios para expressar à sociedade seus posicionamentos, projeções e dados, apresentando um grau de transparência superior àquele exigido pelas normas que regulamentam a sua atuação (Gutierrez, Hennings & Morais, 2009). Os veículos de comunicação do BCB atendem a objetivos distintos. Um desses veículos é o comunicado de decisão de política monetária,¹ o qual apesar de ser bastante considerado pelo mercado ainda é pouco estudado. Pequeno e simples, ele é divulgado logo após a deliberação do Comitê de Política Monetária (COPOM) e se caracteriza por resumir em poucos parágrafos a decisão do Banco Central. Apesar de breve, a nota se constitui em um primeiro indicativo do que se esperar para o próximo passo do ciclo de juros, o que faz com que

*Agradecemos a Carlos Eduardo Gonçalves, Enlison Mattos, Vladimir Teles, Alexandre Cunha, um parecerista anônimo e, especialmente, a Adonias Costa Filho por seus comentários e sugestões. Agradecemos também a Adonias Costa Filho por gentilmente ceder dados sobre sua classificação das atas de reuniões do Banco Central do Brasil. Rodolfo Cabral agradece o apoio financeiro da CAPES e do CNPq. Bernardo Guimaraes agradece o apoio financeiro do CNPq.

[†]Escola de Economia de São Paulo – Fundação Getúlio Vargas (EESP/FGV). Rua Itapeva, 474, Bela Vista, São Paulo-SP. E-mail: rodolfo27@gmail.com

[‡]Escola de Economia de São Paulo – Fundação Getúlio Vargas (EESP/FGV). E-mail: bernardo.guimaraes@fgv.br

¹Nota informativa à imprensa



o mercado crie expectativas e inferências sobre o futuro.²

Este artigo estuda como a indicação de política monetária contida nos comunicados do BCB impacta os mercados. Para isso, classificamos a mensagem de cada nota por meio de um índice semelhante ao proposto por Rosa & Verga (2007). A classificação atribui valores -1, 0 ou 1 no caso de o conteúdo ser, respectivamente, *dovish*, neutro ou *hawkish*. O período compreendido é o intervalo entre setembro de 2003 e dezembro de 2013.

O trabalho traz três resultados principais. O primeiro é que os comunicados das decisões parecem ser bastante eficazes enquanto informantes do futuro da política monetária, influenciando toda a curva de juros e o Ibovespa. No caso do vértice de 360 dias, por exemplo, um anúncio *hawkish* leva a um aumento de cerca de 6 pontos base na taxa de juros, em média, o que equivale, aproximadamente, a uma surpresa positiva de 8.6 pontos base de choque de política monetária.

Como um segundo resultado, encontramos que o comunicado ajuda a explicar muito bem a ata da reunião do COPOM. A ata é publicada na quinta-feira da semana posterior à reunião e tem a proposta de ser mais ampla e analítica na função de esmiuçar as condições que levaram o BCB a tomar determinada decisão. Neste artigo, classificamos a mensagem contida nas atas utilizando a mesma metodologia do comunicado. De modo geral, esse segundo achado nos diz que o comunicado tem papel importante para prever a ata e juntamente com outras variáveis acerta em torno de 92% das vezes a indicação da mesma.

As indicações das atas do COPOM, ao contrário daquelas do comunicado, não parecem impactar os preços de ativos: simples regressões de variação de preço de ativo no valor atribuído à mensagem contida na ata não trazem resultados significativos e expressivos. Uma possível explicação é que o conteúdo da ata já é esperado pelo mercado, sendo a informação antecipada pelo comunicado da semana anterior. A partir dessa evidência, verificamos se os desvios entre os conteúdos realizados e esperados das atas (os “choques de notícias”) conseguem impactar os mercados e encontramos que, de fato, os componentes inesperados das informações afetam a curva de juros. Assim, nosso terceiro achado é que a indicação da ata do COPOM não parece impactar os mercados, mas o choque de notícias da ata afeta a curva de juros. Uma ata *hawkish* quando todos esperavam uma ata neutra tem um impacto de cerca de 11 pontos base nas taxas de juros, em média. Todavia, como há poucas surpresas ao longo dos anos, os resultados não são muito conclusivos.

Os resultados mostram que a maneira como o BCB se expressa é importante e que realmente as palavras parecem influenciar os mercados. No entanto, até a década de 90, pouca atenção havia sido dada à comunicação dos Bancos Centrais. Um exemplo disso é que foi somente a partir de 1999 que o Federal Reserve (FED) passou a divulgar comentários após cada reunião do Federal Open Market Committee (FOMC) contendo alguns balanços sobre inflação e explicitando considerações a respeito das prováveis decisões futuras sobre a política monetária (Costa-Filho, 2008). Recentemente, porém, observa-se uma maior diversificação e ampliação dos canais de informação. Isso porque a política monetária tem se tornado mais prospectiva e a sua transmissão cada vez mais realizada por meio de expectativas (Gutierrez et al., 2009). Por conta disso, gradativamente, ganha consenso a ideia de que a comunicação passou a ser uma ferramenta indispensável da política monetária. A partir desse novo quadro, uma parte considerável da

²Um bom exemplo é visto na nota do Banco Itaú de 09/10/2013 emitida logo após a decisão do COPOM: “O Comitê de Política Monetária do Banco Central do Brasil (COPOM) elevou a taxa de juros básica (Selic) em 0,50 ponto percentual para 9,50% ao ano, como amplamente esperado. A decisão foi unânime. O comunicado que acompanhou a decisão foi similar ao da reunião anterior [...] Em declarações recentes o presidente do Banco Central, Alexandre Tombini, mostrou-se mais confiante com o cenário para a inflação, o que sugeria uma possível mudança na comunicação do COPOM, ocorrida hoje. A manutenção do comunicado, portanto, indica que, na visão do COPOM, o ciclo de alta de juros deve continuar para minimizar as pressões inflacionárias geradas pela depreciação cambial e pelo mercado de trabalho ainda apertado. Dado o crescimento moderado e a estabilidade recente da taxa de câmbio, mantemos por ora nossa expectativa de um ajuste adicional de 0,25 p.p. na reunião de novembro, encerrando o ciclo de aperto monetário com a taxa Selic em 9,75%. No entanto, reconhecemos que a manutenção do comunicado hoje aumenta a probabilidade do ciclo se prolongar para além de 10%, ao passo atual de 0,50 p.p. Esperaremos a ata desta reunião, na próxima quinta-feira, para entender melhor os próximos passos que deverão ser dados pelo Banco Central.” (Goldfajn & Megale, 2013)

literatura tem investigado empiricamente como as palavras podem afetar os mercados.³

Nosso trabalho se baseia na metodologia utilizada em Rosa & Verga (2008), Rosa (2011a), Rosa (2011b) e Rosa (2011c). Nesses artigos, as palavras são classificadas de acordo com um índice proposto em Rosa & Verga (2007). Em síntese, esses trabalhos procuram mostrar como os preços em diversos mercados mudam devido ao componente inesperado da informação das autoridades monetárias, o choque de notícias. Ao fazermos esse tipo de análise para o caso brasileiro, atentamos para as diferenças na estrutura de comunicação do BCB.

Em relação à literatura para o Brasil, Carvalho, Cordeiro & Vargas (2013) é o primeiro trabalho a ter como foco os comunicados do BCB e, até onde sabemos, o único. Os autores analisam os comunicados entre 2007 e 2013 utilizando uma adaptação da metodologia desenvolvida por Lucca & Trebbi (2011). Essa metodologia faz uso de buscas no Google para mensurar o quanto cada informação é percebida como indicativo de aumento ou diminuição de juros. O índice proposto por Lucca & Trebbi é construído a partir da frequência relativa com que sentenças aparecem associadas a dois conceitos opostos, digamos, *hawkish* e *dovish*. Carvalho et al. (2013) constatam que mudanças na linguagem do COPOM parecem ter efeito sobre a curva de juros futuros, mas isso não é válido para o período após o começo do mandato de Alexandre Tombini.

Outros trabalhos estudam principalmente as atas. Costa-Filho & Rocha (2009) encontram evidências de que a comunicação do COPOM é consistente, visto que as palavras contidas nas atas são seguidas por ações na mesma direção. Além disso, baseados em estimativas de regras de Taylor, verificam que o índice construído por eles ajuda a entender a fixação da taxa de juros. O índice considerado por Costa-Filho & Rocha (2009) também é semelhante ao de Rosa & Verga (2007).

Janot & de Souza Mota (2012) mostram que há reduções significativas na volatilidade das taxas de juros futuros após a divulgação das atas do COPOM e dos relatórios de inflação. De maneira complementar, Costa-Filho & Rocha (2010) verificam se o sinal emitido pelo Banco Central para o curso futuro da política monetária é entendido pelo mercado. Encontram que as taxas de juros futuros aumentam no dia de divulgação da ata, o que indicaria um viés conservador da interpretação da comunicação do BCB, ao passo que a volatilidade diminui. Os autores também apontam para o fato de que a reação do mercado independe do conteúdo das atas.

Por fim, Chague, De-Losso, Giovannetti & Manoel (2015) classificam as palavras das atas do COPOM de 2000 a 2012 em alguns grupos semânticos e depois veem a frequência relativa de cada grupo em cada informe. A fim de reduzir a dimensão da matriz resultante da compilação desses dados, eles utilizam o procedimento de análise do componente principal, o que lhes possibilita extrair o “fator de otimismo” de cada ata. Usando a variação desse fator de otimismo, os autores constatam que a comunicação enquanto resumida por esse “fator” consegue mover a curva de juros futuros no dia de divulgação da informação.

Além dessa introdução, esse artigo está organizado da seguinte maneira: a seção 2 aborda quais são as principais características do comunicado, procurando situá-lo na estrutura de comunicação do BCB; a seção 3 expõe os dados utilizados; na seção 4 mostramos os principais resultados encontrados e a metodologia que nos valem; e, por fim, na seção 5 tem-se a conclusão.

2. O COMUNICADO

A adoção do sistema de metas de inflação pelo Brasil ocorreu em junho de 1999 e atribuiu ao Banco Central a responsabilidade pelo cumprimento das metas fixadas utilizando a taxa de juros de curto prazo como instrumento. No regime de metas de inflação, a comunicação exerce um papel importante. No caso do BCB, o conjunto de meios utilizados para se comunicar é amplo e tem como principais veículos de informação as atas das reuniões do COPOM, os comunicados e os relatórios de inflação.

³Há também uma literatura teórica a respeito da comunicação. Ver, por exemplo, Amato, Morris & Shin (2002) e Agranov & Schotter (2013).



As atas do COPOM começaram a ser divulgadas em janeiro de 1998. São publicadas às 8h30min da quinta-feira da semana posterior a cada reunião (prazo estipulado de seis dias úteis). Têm a proposta de serem mais amplas e analíticas na sua função de esmiuçar as condições econômicas que levaram o Banco Central a tomar as decisões e a construir determinadas estratégias. Daí a importância e, por isso, a consideração de que a ata é o veículo de comunicação que recebe maior atenção da mídia e do mercado financeiro (Costa-Filho, 2008). Chama a atenção a manutenção da seção “Implementação da Política Monetária” em que se informam as diretrizes da política do BCB. Também são incluídas no texto algumas avaliações prospectivas da inflação com hipóteses e cenários, mas sem a divulgação de números.

Diferente da ata, o comunicado de decisão de política monetária é simples e curto, apenas alguns parágrafos, prestando-se como uma breve justificativa da decisão tomada e contendo a indicação da votação (desde maio de 2002). Além disso, ele sugere qual o provável ciclo de política monetária e ainda, às vezes, tece considerações sobre a intensidade da política. O comunicado é a primeira indicação do próximo passo do ciclo de política monetária e é divulgado no segundo dia de reunião do COPOM, à noite, quando os mercados já estão fechados.

As reuniões do COPOM acontecem em dois dias. No primeiro, os chefes de departamento apresentam uma vasta análise sobre diversos temas. No segundo, os diretores de Política Monetária e de Política Econômica propõem alternativas para a taxa de juros de curto prazo. Após ponderações, procede-se à votação. O resultado é imediatamente divulgado à imprensa ao mesmo tempo em que a nota é expedida. A partir de 2000 as reuniões se tornaram mensais, fato que vigorou até 2006 quando o número de reuniões passou a ser oito ao ano (BCB, 2012).

Uma análise do comunicado da maneira que propomos aqui é mais factível a partir de 2003, pois apesar de existir desde 2000 para informar a decisão do COPOM, ele foi mudando ao longo do tempo e, aos poucos, incorporando a função de ser uma nota com mais significado. Por ser sucinto e simples, com frequência se consegue compreender o que o BCB quis dizer a partir de uma única observação. Desse modo, se o BCB repete a mesma nota em uma outra divulgação, fica fácil saber como o recado deve ser lido.

A Tabela 1 compara o comunicado do BCB com os de outros bancos centrais.⁴ Todos os bancos centrais da amostra publicam algum texto após as decisões e, como particularidade do caso brasileiro, o comunicado do BCB parece ser o menor dentre aqueles aqui comparados. Apesar disso, o comunicado do COPOM contém a votação do comitê, algo que não é compartilhado por todos. Ademais, o Brasil parece ser mais exceção do que regra ao não divulgar cenários e balanços de riscos imediatamente após a decisão de política monetária.

3. DADOS

Nossa amostra consiste em atas e comunicados de setembro de 2003 a dezembro de 2013, totalizando 90 observações de cada um.⁵ Para ambos os veículos foram construídos índices por meio de interpretação, tendo em vista a “tradução” das palavras em números. Sendo assim, buscou-se classificar os instrumentos de comunicação em algum grupo; -1, 0, 1; em que -1 indica uma compreensão de que o tom das palavras sugere redução futura da taxa de juros básica; 0 sinaliza que as notícias são neutras, não apontam nenhuma tendência senão a manutenção; e 1 indica que o Banco Central demonstra uma maior intenção em subir os juros no futuro. Essa metodologia é semelhante àquela empregada por Rosa & Verga (2007).

⁴Tabela retirada da apresentação de Gutierrez et al. (2009) em que podemos observar as semelhanças e diferenças entre os comunicados/notas das autoridades monetárias. As autoras analisam os bancos centrais de 12 regiões. São elas: Austrália, Brasil, Canadá, Chile, Israel, Nova Zelândia, Reino Unido, Suécia, Japão, Suíça, União Europeia e EUA.

⁵Para esse período, as únicas reuniões em que não consideramos as informações são as de março e de abril de 2004. Para a primeira, na base de dados original, Adonias Costa Filho não classifica a ata desta reunião. No segundo caso, o encontro de abril, a reunião foi encerrada durante a tarde, por isso optou-se por excluí-la.

Tabela 1. Comunicados das decisões.

	Comunicados					
	Divulgação	Dimensão	Votação	Cenário	Balanco de Riscos	Perspectivas
Austrália	Sim	Médio	Não	Sim	Não	Não
Brasil	Sim	Curto	Sim	Não	Não	Não
Canadá	Sim	Médio	Consenso	Sim	Sim	Sim
Chile	Sim	Médio	Sim	Sim	Sim	Sim
Israel	Sim	Longo	N.A.	Sim	Sim	Não
N. Zelândia	Sim	Médio	N.A.	Sim	Sim	Sim
Reino Unido	Sim	Médio	Não	Sim	Sim	Não
Suécia	Sim	Longo	Não	Sim	Sim	Sim
Japão	Sim	Médio	Sim	Não	Não	Não
Suíça	Sim	M/L	Não	Sim	Sim	Sim
ECB	Sim	Longo	Não	Sim	Sim	Sim
EUA	Sim	Médio	Sim	Sim	Sim	Sim

Por ser um entendimento do que se quis dizer, a classificação utilizada é bastante passível de erros. Contudo, interpretações errôneas também podem ter sido adotadas pelo mercado à época das divulgações, em que “errônea” aqui significa: diferente do que foi desejado pelo BCB. Assim, o ganho desse tipo de análise também é uma de suas principais debilidades: a subjetividade. No caso do comunicado, os mesmos são simples e diretos, o que de certa maneira ajuda a reduzir tal problema.

Para as atas das reuniões do COPOM entre 2003 e 2006, a classificação foi feita por Adonias Costa Filho em virtude de seus trabalhos sobre comunicação (Costa-Filho, 2008; Costa-Filho & Rocha, 2009, 2010). Para o período posterior a 2006, a análise é baseada em um glossário elaborado por Adonias Costa Filho que codifica uma série de expressões e palavras indicadoras dos posicionamentos do BCB (Costa-Filho, 2008, p.18; Costa-Filho & Rocha, 2009, p.409). Exemplos de atas classificadas podem ser encontrados em Costa-Filho (2008, pp.57–60). No Apêndice encontra-se uma tabela com os valores dos comunicados e das atas em cada uma das reuniões analisadas.

Para os comunicados, a análise semântica é complementada pelo entendimento das notas precedentes. Tomemos os exemplos das reuniões 169^a, 170^a, 171^a, 172^a, e 173^a. Depois de sequenciais quedas na taxa Selic, na 169^a reunião o COPOM novamente reduziu os juros e alertou que um possível ajuste adicional ainda era cabível, dando um tom de continuação das reduções da Selic. Por conta disso, o comunicado tem a classificação de -1 (*dovish*). Vejamos:

O Copom decidiu, por unanimidade, reduzir a taxa Selic para 7,50% a.a., sem viés. Considerando os efeitos cumulativos e defasados das ações de política implementadas até o momento, que em parte se refletem na recuperação em curso da atividade econômica, o Copom entende que, se o cenário prospectivo vier a comportar um ajuste adicional nas condições monetárias, esse movimento deverá ser conduzido com máxima parcimônia. Votaram pela redução da taxa Selic para 7,50% os seguintes membros do Comitê: Alexandre Antonio Tombini, Presidente, Aldo Luiz Mendes, Altamir Lopes, Anthero de Moraes Meirelles, Carlos Hamilton Vasconcelos Araújo, Luiz Awazu Pereira da Silva, Luiz Edson Feltrim e Sidnei Corrêa Marques.

De fato, o que aconteceu na reunião que se seguiu foi uma nova queda dos juros, corroborando com a indicativa acima (de 7,5% a.a para 7,25% a.a). A seguir, temos a justificativa para tal queda na 170^a reunião:



O Copom decidiu reduzir a taxa Selic para 7,25% a.a., sem viés, por 5 votos a favor e 3 votos pela manutenção da taxa Selic em 7,50% a.a. Considerando o balanço de riscos para a inflação, a recuperação da atividade doméstica e a complexidade que envolve o ambiente internacional, o Comitê entende que a estabilidade das condições monetárias por um período de tempo suficientemente prolongado é a estratégia mais adequada para garantir a convergência da inflação para a meta, ainda que de forma não linear. Votaram pela redução da taxa Selic para 7,25% a.a. os seguintes membros do Comitê: Alexandre Antonio Tombini, Presidente, Aldo Luiz Mendes, Altamir Lopes, Luiz Awazu Pereira da Silva e Luiz Edson Feltrim. Votaram pela manutenção da taxa Selic em 7,50% a.a. os seguintes membros do Comitê: Anthero de Moraes Meirelles, Carlos Hamilton Vasconcelos Araújo e Sidnei Corrêa Marques.

A redução não ter sido consensual é um indicador de que o cenário havia mudado e de que o ritmo de queda poderia ser revisto. Aliada a isso, a frase “o Comitê entende que a estabilidade das condições monetárias por um período de tempo suficientemente prolongado é a estratégia mais adequada para garantir a convergência da inflação para a meta” é bastante indicativa de que o ciclo de reduções parecia ter chegado ao fim e de que ajustes adicionais pareciam ser inoportunos. Assim, esse comunicado corresponde a um índice 0, ou seja, neutro. A 171^a reunião concretizou tal perspectiva de manutenção dos juros e novamente sinalizou para a continuação da Selic no patamar de 7,25% a.a. Vejamos:

O Copom decidiu, por unanimidade, manter a taxa Selic em 7,25% a.a., sem viés. Considerando o balanço de riscos para a inflação, a recuperação da atividade doméstica e a complexidade que envolve o ambiente internacional, o Comitê entende que a estabilidade das condições monetárias por um período de tempo suficientemente prolongado é a estratégia mais adequada para garantir a convergência da inflação para a meta, ainda que de forma não linear. Votaram por essa decisão os seguintes membros do Comitê: Alexandre Antonio Tombini, Presidente, Aldo Luiz Mendes, Altamir Lopes, Anthero de Moraes Meirelles, Carlos Hamilton Vasconcelos Araújo, Luiz Awazu Pereira da Silva, Luiz Edson Feltrim e Sidnei Corrêa Marques.

É notável o uso da frase “o Comitê entende que a estabilidade das condições monetárias por um período de tempo suficientemente prolongado é a estratégia mais adequada para garantir a convergência da inflação para a meta”, novamente. Sendo assim, tal comunicado é classificado como neutro, 0. Para a reunião 172^a, a Selic foi mantida em 7,25%. Segue o comunicado:

O Copom decidiu, por unanimidade, manter a taxa Selic em 7,25% a.a., sem viés. Considerando o balanço de riscos para a inflação, que apresentou piora no curto prazo, a recuperação da atividade doméstica, menos intensa do que o esperado, e a complexidade que ainda envolve o ambiente internacional, o Comitê entende que a estabilidade das condições monetárias por um período de tempo suficientemente prolongado é a estratégia mais adequada para garantir a convergência da inflação para a meta. Votaram por essa decisão os seguintes membros do Comitê: Alexandre Antonio Tombini (Presidente), Aldo Luiz Mendes, Altamir Lopes, Anthero de Moraes Meirelles, Carlos Hamilton Vasconcelos Araújo, Luiz Awazu Pereira da Silva, Luiz Edson Feltrim e Sidnei Corrêa Marques.

Com essa nota vemos que a frase que caracterizou os dois comunicados anteriores ainda está presente dando um caráter neutro à informação, 0. Todavia, percebe-se que o comitê fez questão de enfatizar a piora no balanço de riscos para a inflação, o que é um vetor extremamente importante para um aumento de juros. Vê-se também as preocupações com a atividade doméstica e o ambiente internacional, os quais parecem ter tido bastante relevância na decisão e podem ter justificado o uso da “frase neutra”. Por fim, consideremos o comunicado da 173^a reunião:

Avaliando a conjuntura macroeconômica e as perspectivas para a inflação, o Copom decidiu, por unanimidade, manter a taxa Selic em 7,25% a.a., sem viés. O Comitê irá acompanhar a evolução do cenário macroeconômico até sua próxima reunião, para então definir os próximos passos na sua estratégia de política monetária. Votaram por essa decisão os seguintes membros do Comitê: Alexandre Antonio Tombini (Presidente), Aldo Luiz Mendes, Altamir Lopes, Anthero de Moraes Meirelles, Carlos Hamilton Vasconcelos Araújo, Luiz Awazu Pereira da Silva, Luiz Edson Feltrim e Sidnei Corrêa Marques.

Esse comunicado traz uma clara mudança de tom do comitê. A estrutura semelhante dos últimos deu espaço a uma posição de cautela e maior foco à inflação, a qual tínhamos visto que era um vetor de aumento de juros, dada a piora no balanço de riscos. Assim, esse comunicado recebe o valor de +1 (*hawkish*).⁶ Na reunião que se seguiu, o COPOM subiu os juros, corroborando com a informação.

Os outros dados que utilizamos nesse trabalho são provenientes de diversas fontes. A série de inflação mensal do IPCA e os valores diários⁷ do *Emerging Markets Bond Index*, o EMBI calculado pelo banco J.P. Morgan, foram retirados do Ipeadata.⁸ Os números para a meta da taxa Selic, para câmbio, dólar americano, e para as maturidades até 360 dias dos Swaps Pré-DI, são oriundos do sítio do BCB.⁹ Para as maturidades entre 1,5 ano e 5 anos, para o Ibovespa e também para a expectativa de Selic imediatamente antes das reuniões,¹⁰ utilizamos a base de dados da Bloomberg.

A Tabela 2 traz um quadro síntese com as principais estatísticas descritivas dos ativos estudados.

4. RESULTADOS

4.1. O comunicado afeta os mercados

Na divisão utilizando a metodologia de classificação proposta em Rosa & Verga (2007), temos que os 90 comunicados se distribuem bem entre as três opções de tonalidade semântica: como -1 temos 29 observações, com uma interpretação neutra há 29, e os comunicados que apresentam a retórica de aumento de juros no futuro somam 32 (Figura 1).

Nosso interesse reside em saber se o conteúdo dos comunicados tem impacto sobre os preços. Assim, neste primeiro exercício, verificamos se o retorno de um ativo k varia com o índice. Como o comunicado é divulgado na noite do segundo dia de reunião do COPOM, t , consideramos $r_{k,t}$ como o preço no dia do anúncio e $r_{k,t+1}$ como aquele que vigorou no dia seguinte, $t+1$. Dado que este veículo é, por definição, a notícia de decisão de política monetária, controlamos também as nossas regressões pelo que chamamos

⁶O mercado também parece ter interpretado dessa forma.

O mercado de juros futuros viveu um dia de fortíssimo volume de negócios e intensos ajustes das taxas, refletindo as mudanças das apostas provocadas pelo comunicado do Comitê de Política Monetária (Copom) do Banco Central de ontem, que manteve a Selic em 7,25% ao ano. Tudo indica que o volume do dia será recorde: somente o contrato de Depósito Interfinanceiro (DI) de julho de 2013 movimentou 1,5 milhão de contratos hoje. Já o DI janeiro/2014 registrou 1,2 milhão de contratos. Segundo operadores, investidores ampliaram as fichas na ideia de que o ciclo de aperto monetário será mais prolongado depois da leitura do comunicado da reunião de ontem. Segundo cálculos de operadores, os DIs com vencimentos mais próximos indicam chance de alta de 0,25 ponto da Selic em abril. Para as reuniões de maio, julho e agosto, os DIs já precificam altas em doses superiores a 0,30 ponto, e mais alguma alta residual nos encontros seguintes. A leitura do mercado é que o Banco Central abriu a porta para começar a subir os juros, mas não deixou claro que isso deva ocorrer já em abril. O mercado mantém, de todo modo, a aposta em uma alta já no mês que vem, embora parte dos analistas considere que cresce a chance de o ajuste começar apenas em maio. Isso por causa do comunicado utilizado ontem, já empregado em outras situações em que o aumento de juros veio apenas duas reuniões depois. [...] (Pinto, 2013)

⁷Utilizados em porcentagem.

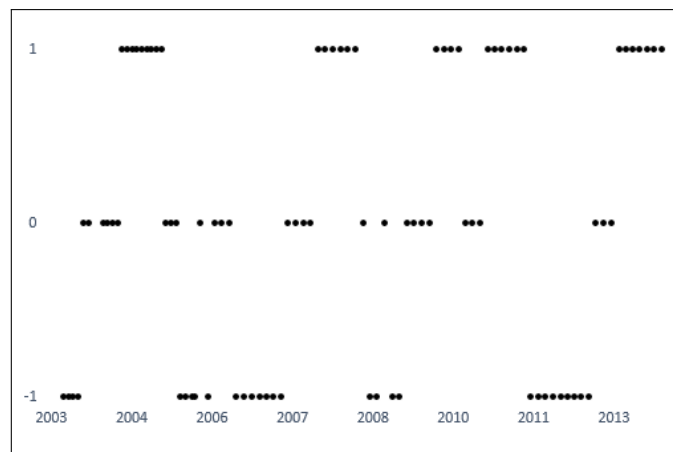
⁸<http://www.ipeadata.gov.br>

⁹<http://www.bcb.gov.br>

¹⁰Pesquisa Bloomberg com os analistas de mercado.

**Tabela 2.** Estatísticas Descritivas (01/09/2003 até 23/12/2013).

	30d	60d	90d	120d	180d	360d	540d
Média	12,48	12,47	12,47	12,48	12,51	12,65	12,83
Mediana	11,74	11,76	11,82	11,77	11,79	11,99	12,22
Desv. Padrão	3,65	3,62	3,59	3,56	3,50	3,34	3,18
Mín	6,93	6,96	6,97	6,99	6,92	6,86	6,99
Máx	21,05	20,46	20,02	19,88	19,74	19,42	19,69
Observações	2.565	2.565	2.564	2.564	2.565	2.565	2.545
	2anos	3anos	4anos	5anos	câmbio	ibov	
Média	12,98	13,16	13,28	13,36	2,11	48.028,89	
Mediana	12,34	12,40	12,49	12,50	2,04	52.996,78	
Desv. Padrão	3,05	2,91	2,87	2,90	0,40	16.084,31	
Mín	7,31	7,89	8,27	8,51	1,53	15.352,12	
Máx	20,41	21,41	22,42	23,03	3,21	73.516,81	
Observações	2.545	2.544	2.544	2.544	2.595	2.556	

Figura 1. Índice do comunicado.

de choque de política monetária, CPM . O choque de política monetária em t nada mais é do que o desvio do valor da meta da taxa Selic adotado pelas autoridades monetárias, r_t^* , para com a média das expectativas dos analistas, r_t^e :

$$CPM_t = r_t^* - r_t^e,$$

e sua trajetória é apresentada na Figura 2.

Dessa forma, a regressão que estimamos é

$$r_{k,t+1} - r_{k,t} = \alpha + \beta_{IndiceC} IndiceC_t + \beta_{CPM} CPM_t + \beta_X X_t + \epsilon_{t+1},$$

em que $IndiceC_t$ é o índice do comunicado e X_t é um vetor de variáveis de controle.¹¹

Na Tabela 3, observamos que o índice do comunicado é significativo estatisticamente para todas as maturidades até 5 anos e também para o Ibovespa, apresentando coeficientes positivos. Vemos, portanto,

¹¹Nos resultados reportados, por concisão, há apenas uma variável de controle, a medida de risco país dada pelo EMBI. Quando possível, incluir o Ibovespa e o câmbio como variáveis independentes, por exemplo, não altera os resultados de interesse.

Figura 2. Choque de política monetária.

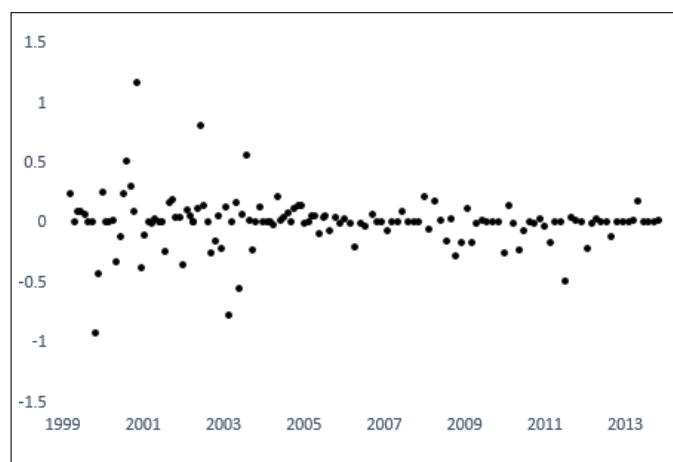


Tabela 3. Indicação de Política Monetária – índice do comunicado.

	(1) 30d	(2) 60d	(3) 90d	(4) 120d	(5) 180d	(6) 360d	(7) 540d
Índice	0.0225*** (0.00740)	0.0213** (0.00869)	0.0379*** (0.0113)	0.0469*** (0.0134)	0.0530*** (0.0156)	0.0595*** (0.0189)	0.0609*** (0.0211)
CPM	0.758*** (0.0453)	0.785*** (0.0532)	0.821*** (0.0691)	0.855*** (0.0820)	0.864*** (0.0951)	0.695*** (0.116)	0.551*** (0.129)
EMBI	0.00600 (0.00417)	0.0110** (0.00490)	0.0174*** (0.00637)	0.0194** (0.00755)	0.0234*** (0.00876)	0.0350*** (0.0107)	0.0332*** (0.0119)
Constante	-0.0258* (0.0139)	-0.0368** (0.0164)	-0.0560** (0.0213)	-0.0602** (0.0253)	-0.0681** (0.0293)	-0.0988*** (0.0357)	-0.0855** (0.0398)
Observações	90	90	90	90	90	90	90
R-quadrado	0.801	0.760	0.699	0.652	0.601	0.460	0.340
	(8) 2anos	(9) 3anos	(10) 4anos	(11) 5anos	(12) câmbio	(13) ibov	
Índice	0.0551** (0.0221)	0.0450** (0.0213)	0.0416* (0.0221)	0.0384* (0.0223)	0.00110 (0.00293)	216.9** (107.7)	
CPM	0.464*** (0.135)	0.329** (0.130)	0.219 (0.135)	0.145 (0.136)	0.000739 (0.0179)	-1,827*** (658.5)	
EMBI	0.0292** (0.0124)	0.0199 (0.0120)	0.0255** (0.0124)	0.0215* (0.0126)	-0.00179 (0.00165)	-68.76 (60.67)	
Constante	-0.0656 (0.0416)	-0.0356 (0.0401)	-0.0454 (0.0416)	-0.0325 (0.0420)	0.00139 (0.00551)	435.7** (202.9)	
Observações	90	90	90	90	90	90	
R-quadrado	0.261	0.170	0.128	0.088	0.016	0.124	

Erro padrão entre parênteses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$



que o comunicado como indicador do futuro da política monetária parece ser importante para explicar a variação de preço dos ativos. Os coeficientes observados devem ser lidos como um aumento nos preços se a autoridade monetária comunicar um tom mais *hawkish* em comparação a palavras mais neutras que indicam a manutenção da taxa de juros (o simétrico vale para um anúncio *dovish*). Nesse sentido, por exemplo, no caso do vértice de 360 dias, se o anúncio for *hawkish*, os juros aumentam, em média, cerca de 6 pontos base, o que equivale, aproximadamente, a uma surpresa positiva de 8,6 pontos base de choque de política monetária.

Com relação ao Ibovespa, percebemos que o choque de política monetária e o indicador possuem sinais opostos. Uma hipótese plausível para isso é que o índice mais do que somente sugerir o possível curso dos juros no futuro, transmite também características do Banco Central, como o seu comprometimento, por exemplo. Talvez o mercado de ações interprete essa informação dessa maneira. Isso é diferente do choque de política monetária, que é a realização da mudança dos juros em si.

4.2. O comunicado ajuda a prever a ata

Para as atas do COPOM, verificamos que do total de 90 observações, 33 são classificadas com um tom prospectivo de redução dos juros, 24 parecem indicar a manutenção da Selic, e outras 33 indicam aumento de juros no futuro (Figura 3).

Nesta seção, estudamos a interação entre comunicados e atas do COPOM. Dado que a indicação do comunicado consegue impactar os mercados, cabe questionar se o comunicado antecipa a mensagem contida na ata. Por refletirem a visão do mesmo comitê, é de se esperar que para uma mesma reunião, as mensagens de ambos os veículos sejam semelhantes.

Estimamos a relação entre o comunicado e a ata, no dia de divulgação dessa última, t , utilizando um Probit ordenado de acordo com a seguinte equação:

$$IndiceA_t = \gamma_1 S_t + \gamma_2 IndiceC_t + \epsilon_t.$$

A variável dependente é o índice para a ata, $IndiceA_t$. O valor atribuído ao comunicado da mesma reunião, $IndiceC_t$, entra como variável explicativa, assim como o vetor S_t composto pelo desvio da inflação acumulada 12 meses com relação à meta de inflação e a inclinação da curva de juros no momento imediatamente anterior à declaração, medida pela diferença entre as taxas de Swaps Pré-DI de 90 e 30 dias.

Como mostra a Tabela 4, todas as variáveis são significativas aos níveis usuais. Com relação à capacidade de previsão da regressão, o Pseudo- R^2 é de 0,847. O comunicado ajuda bastante a explicar a

Figura 3. Índice da ata.

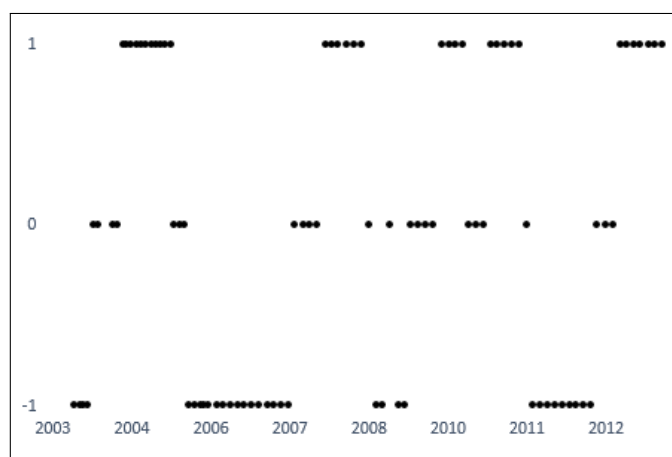


Tabela 4. Probit ordenado para a mensagem da ata.

	Índice	Erro
Comunicado	3.124***	(0.616)
Inclinação	5.805***	(1.790)
Inflação	0.445*	(0.260)
Limite 1	-1.526***	(0.468)
Limite 2	2.854***	(0.789)
Observações	90	
Pseudo R-quadrado	0.847	
Log Verossimilhança	-14.969	

Erro padrão entre parênteses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

ata, pois quando repetimos o mesmo exercício sem o incluir, o modelo é bastante prejudicado, com um Pseudo- R^2 de 0,583.

Para o índice de acertos, a porcentagem é consideravelmente alta com o comunicado, como mostra a Tabela 5.

Tabela 5. Avaliação da previsão sobre a ata.

Índice	Obs.	Correto	Incorreto	% Correto	% Incorreto
-1	33	30	3	90.909	9.091
0	24	22	2	91.667	8.333
1	33	31	2	93.939	6.061
Total	90	83	7	92.222	7.778

4.3. As poucas surpresas da ata parecem afetar o mercado

De posse do índice para as atas, um passo natural é verificar se a ata traz consigo alguma novidade, uma nova informação, ou seja, se o conteúdo das atas consegue influenciar os mercados. Em um primeiro momento, a regressão utilizada para estudarmos tais impactos é idêntica à do comunicado. Como esse veículo é divulgado na manhã do sexto dia útil após a reunião do comitê, consideramos $r_{k,t-1}$ como o preço no dia anterior à ata, $t-1$, e $r_{k,t}$ como aquele que vigorou no dia da divulgação, t . Assim,¹²

$$r_{k,t} - r_{k,t-1} = \alpha + \beta_{IndiceA} IndiceA_t + \epsilon_t. \quad (1)$$

De modo geral, encontramos que as indicações das atas não parecem ter muita influência sobre os preços dos ativos,¹³ como mostra a Tabela 6.

Uma possível explicação para esse fato é que a indicação de política monetária já é conhecida no dia de divulgação da meta da Selic, papel desempenhado pelo comunicado. Diante da evidência acima, buscamos averiguar, então, se as surpresas contidas nas atas conseguem impactar os mercados. Para

¹²Como a ata é expedida somente na semana seguinte à decisão do COPOM, não faz sentido incluirmos o choque de política monetária como controle na regressão.

¹³Os únicos ativos para os quais o índice foi significativo aos níveis usuais foram o vértice de 30 dias e o câmbio. No entanto, os coeficientes são pequenos.

**Tabela 6.** Indicação de Política Monetária – índice ata.

	(1) 30d	(2) 60d	(3) 90d	(4) 120d	(5) 180d	(6) 360d	(7) 540d
Índice	0.00409* (0.00242)	-0.000303 (0.00420)	0.00848 (0.00627)	0.00833 (0.00840)	0.00803 (0.0112)	0.00636 (0.0156)	0.00915 (0.0177)
Constante	0.00233 (0.00207)	0.000111 (0.00360)	-0.000556 (0.00537)	0.00289 (0.00719)	0.00611 (0.00960)	0.00567 (0.0134)	0.00814 (0.0152)
Observações	90	90	90	90	90	90	90
R-quadrado	0.032	0.000	0.020	0.011	0.006	0.002	0.003
	(8) 2anos	(9) 3anos	(10) 4anos	(11) 5anos	(12) câmbio	(13) ibov	
Índice	0.0127 (0.0187)	0.0175 (0.0199)	0.0280 (0.0218)	0.0260 (0.0234)	0.00453** (0.00208)	-125.9 (114.3)	
Constante	0.0120 (0.0160)	0.00813 (0.0170)	0.0141 (0.0187)	0.0151 (0.0201)	0.000219 (0.00178)	-53.41 (97.84)	
Observações	90	90	90	90	90	90	
R-quadrado	0.005	0.009	0.018	0.014	0.051	0.014	

Erro padrão entre parênteses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

isso, inicialmente, utilizamos as probabilidades estimadas e calculamos as diferenças entre os conteúdos realizados e esperados das atas, o que denominamos de choques de notícias, *CNA*:

$$CNA_t = ÍndiceA_t - E_t[ÍndiceA_t],$$

onde o valor esperado da comunicação é dado por

$$E_t[ÍndiceA_t] = \sum_{i=-1}^1 \Pr(ÍndiceA_t = i) \cdot i.$$

A Figura 4 mostra os valores para os choques de notícias proporcionados pela ata. Uma consequência de prever bem a ata futura é que os choques de notícias apresentam muitos valores iguais a zero ou próximos disso, o que torna difícil identificar o impacto de surpresas contidas na ata sobre o mercado.

Nosso interesse, nesta etapa, consiste em verificar o impacto do coeficiente β_{CNA} na seguinte equação:

$$r_{k,t} - r_{k,t-1} = \alpha + \beta_{CNA} CNA_t + \epsilon_t.$$

Como resultado, encontramos que tendo o comunicado como predictor da ata do COPOM, o choque de notícias da ata parece mover a curva de juros. De acordo com a Tabela 7, as maturidades mais longas são as que mais respondem ao choque. Em média, o efeito de um choque de dimensão 1 (que corresponde a uma ata *hawkish* quando todos esperavam uma ata neutra) é de, aproximadamente, 11 pontos base. Para o Ibovespa e câmbio, não achamos nenhum efeito significativo. Devido ao pequeno número de surpresas nas atas, esses resultados são menos conclusivos que os anteriores.

5. CONCLUSÃO

O comunicado da decisão do COPOM parece ser bastante eficaz enquanto informante do futuro da política monetária, influenciando toda a curva de juros e o Ibovespa. O comunicado ajuda também a prever o conteúdo da ata de reunião do Banco Central. Há muito poucas surpresas quando a ata é divulgada. Portanto, o comunicado do BCB ocupa papel central na comunicação de política monetária no Brasil.

Figura 4. Choque de notícias.

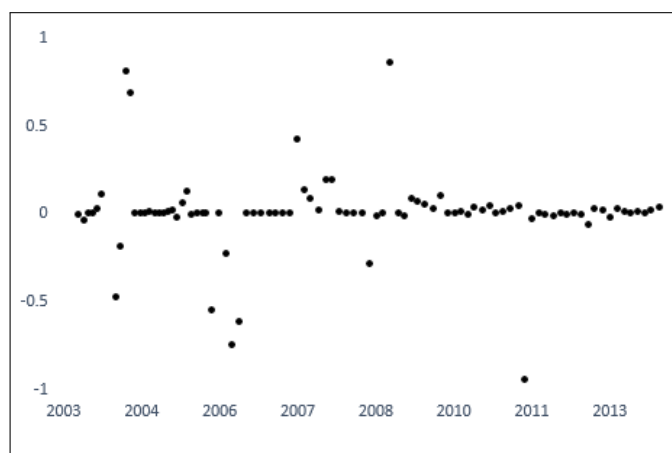


Tabela 7. Choque de notícias – ata.

	(1) 30d	(2) 60d	(3) 90d	(4) 120d	(5) 180d	(6) 360d	(7) 540d
Choque de notícias	0.0172*	0.0260*	0.0385	0.0518*	0.0710*	0.108*	0.122*
	(0.00896)	(0.0154)	(0.0232)	(0.0310)	(0.0412)	(0.0572)	(0.0649)
Constante	0.00234	0.000128	-0.000530	0.00292	0.00616	0.00574	0.00823
	(0.00206)	(0.00354)	(0.00534)	(0.00712)	(0.00947)	(0.0131)	(0.0149)
Observações	90	90	90	90	90	90	90
R-quadrado	0.040	0.031	0.030	0.031	0.033	0.039	0.039
	(8) 2anos	(9) 3anos	(10) 4anos	(11) 5anos	(12) câmbio	(13) ibov	
Choque de notícias	0.151**	0.156**	0.187**	0.232***	-0.00886	-394.3	
	(0.0680)	(0.0726)	(0.0795)	(0.0844)	(0.00791)	(426.6)	
Constante	0.0121	0.00824	0.0142	0.0153	0.000213	-53.67	
	(0.0156)	(0.0167)	(0.0183)	(0.0194)	(0.00182)	(98.04)	
Observações	90	90	90	90	90	90	
R-quadrado	0.053	0.050	0.059	0.079	0.014	0.010	

Erro padrão entre parênteses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

REFERÊNCIAS BIBLIOGRÁFICAS

- Agranov, M., & Schotter, A. (2013). Language and government coordination: An experimental study of communication in the announcement game. *Journal of Public Economics*, 104, 26–39.
- Amato, J. D., Morris, S. & Shin, H. S. (2002). Communication and monetary policy. *Oxford Review of Economic Policy*, 18(4), 495–503.
- Banco Central do Brasil (BCB). (2012). *Copom: Definição e histórico*. Disponível em: <http://www.bcb.gov.br/>
- Carvalho, C., Cordeiro, F. & Vargas, J. (2013). Just words? A quantitative analysis of the communication of the Central Bank of Brazil. *Revista Brasileira de Economia*, 67(4), 443–455.
- Chague, F. D., De-Losso, R., Giovannetti, B. C. & Manoel, P. (2015). Central bank communication affects long-term interest rates. *Revista Brasileira de Economia*, 69(2), 147–162.
- Costa-Filho, A. E. (2008). *Um estudo sobre o papel da comunicação na política monetária* (Dissertação de Mes-

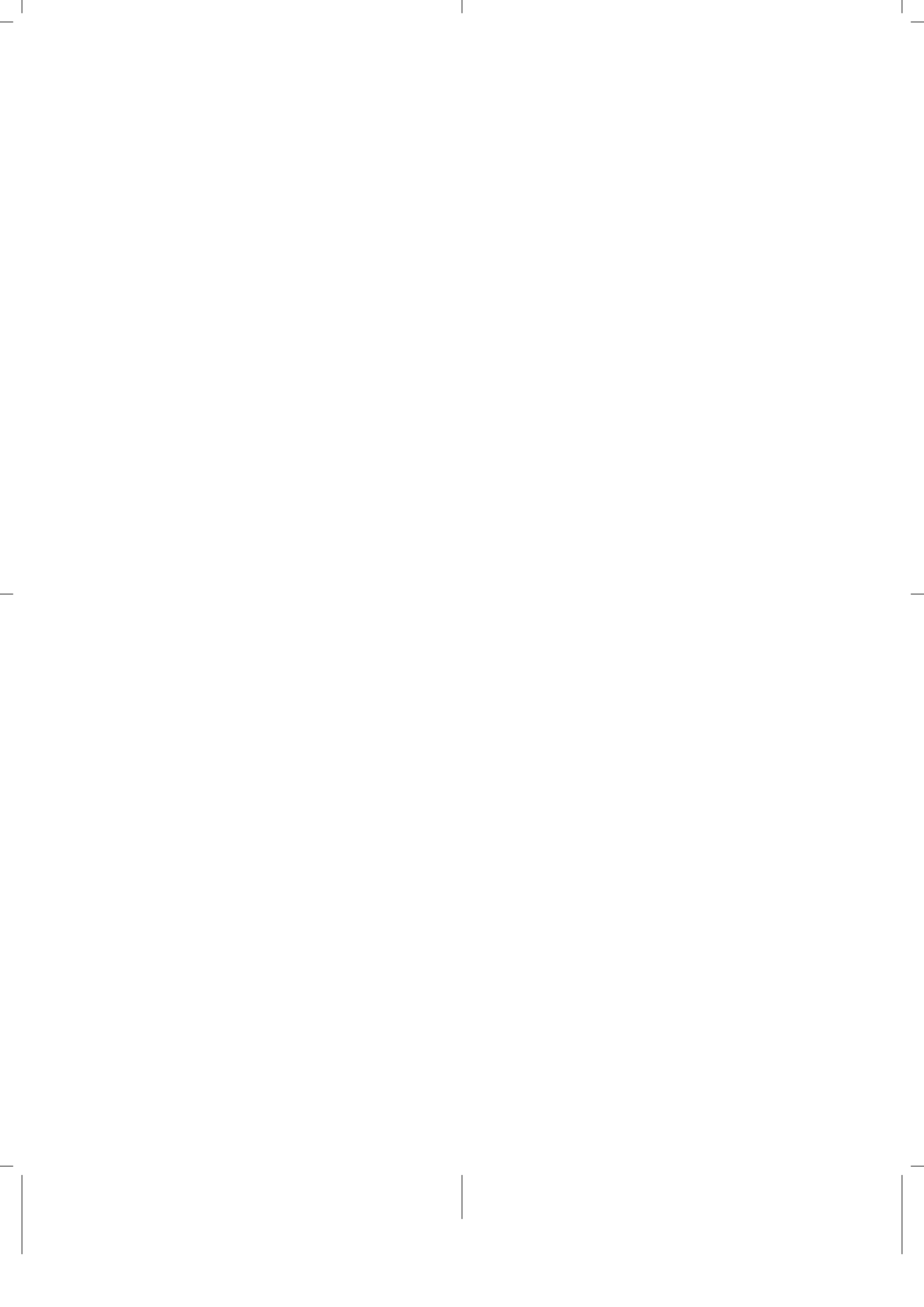


- trado em Economia, Universidade de São Paulo, FEA-USP). Disponível em: <http://www.teses.usp.br/teses/disponiveis/12/12138/tde-04062008-110358/publico/Adonias.pdf>
- Costa-Filho, A. E., & Rocha, F. (2009). Comunicação e Política Monetária no Brasil. *Revista Brasileira de Economia*, 63(4), 405–422.
- Costa-Filho, A. E., & Rocha, F. (2010). Como o mercado de juros futuros reage à comunicação do Banco Central? *Economia Aplicada*, 14(3), 265–292.
- Goldfajn, I., & Megale, C. (2013). *Copom mantém o passo*. Acessado em 24/04/2015: <https://www.itau.com.br/itaubba-pt/analises-economicas/publicacoes/macro-brasil/copom-mantem-o-passo>
- Gutierrez, M., Hennings, K. & Morais, A. (2009, maio 14). Estratégia de comunicação da política monetária no regime de metas para a inflação [Slides]. In *XI Seminário Anual de Metas para a Inflação — Banco Central do Brasil*, Rio de Janeiro. Disponível em: http://www.bcb.gov.br/Pec/Depep/Seminarios/2009_XISemAnualMetasInflBCB/Arquivos/2009_XISemAnualMetasInflBCB_KaterineHennings_pt.pdf
- Janot, M., & de Souza Mota, D. E.-J. (2012, jan). *O impacto da comunicação do Banco Central do Brasil sobre o mercado financeiro* (Trabalhos para Discussão N° 265). Brasília: Banco Central do Brasil. Disponível em: <http://www.bcb.gov.br/pec/wps/port/TD265.pdf>
- Lucca, D. O., & Trebbi, F. (2011). *Measuring central bank communication: An automated approach with application to FOMC statements*. Disponível em: <http://faculty.arts.ubc.ca/ftrebbi/research/lt.pdf>
- Pinto, L. (2013, março 07). Juros reforçam aposta em alta da Selic após comunicado do COPOM. *Valor Econômico*. Disponível em: <http://www.valor.com.br/financas/3036434/juros-reforcam-aposta-em-alta-da-selic-apos-comunicado-do-copom>
- Rosa, C. (2011a). The high-frequency response of exchange rates to monetary policy actions and statements. *Journal of Banking & Finance*, 35(2), 478–489.
- Rosa, C. (2011b). Talking less and moving the market more: Evidence from the ECB and the FED. *Scottish Journal of Political Economy*, 58(1), 51–81.
- Rosa, C. (2011c). Words that shake traders: The stock market's reaction to central bank communication in real time. *Journal of Empirical Finance*, 18(5), 915–934.
- 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), 146–175.
- Rosa, C., & Verga, G. (2008). The impact of central bank announcements on asset prices in real time. *International Journal of Central Banking*, 4(2), 175–217.

APÊNDICE. VALORES DE COMUNICADOS E ATAS

Tabela A-1. Valores de comunicados e atas.

Reunião	Comunicado	Ata	Reunião	Comunicado	Ata
88 ^a	-1	-1	135 ^a	1	1
89 ^a	-1	-1	136 ^a	1	1
90 ^a	-1	-1	137 ^a	1	1
91 ^a	-1	-1	138 ^a	0	0
92 ^a	0	0	139 ^a	-1	-1
93 ^a	0	0	140 ^a	-1	-1
96 ^a	0	0	141 ^a	0	0
97 ^a	0	0	142 ^a	-1	-1
98 ^a	0	1	143 ^a	-1	-1
99 ^a	0	1	144 ^a	0	0
100 ^a	1	1	145 ^a	0	0
101 ^a	1	1	146 ^a	0	0
102 ^a	1	1	147 ^a	0	0
103 ^a	1	1	148 ^a	1	1
104 ^a	1	1	149 ^a	1	1
105 ^a	1	1	150 ^a	1	1
106 ^a	1	1	151 ^a	1	1
107 ^a	1	1	152 ^a	0	0
108 ^a	1	1	153 ^a	0	0
109 ^a	0	0	154 ^a	0	0
110 ^a	0	0	155 ^a	1	1
111 ^a	0	0	156 ^a	1	1
112 ^a	-1	-1	157 ^a	1	1
113 ^a	-1	-1	158 ^a	1	1
114 ^a	-1	-1	159 ^a	1	1
115 ^a	-1	-1	160 ^a	1	0
116 ^a	0	-1	161 ^a	-1	-1
117 ^a	-1	-1	162 ^a	-1	-1
118 ^a	0	-1	163 ^a	-1	-1
119 ^a	0	-1	164 ^a	-1	-1
120 ^a	0	-1	165 ^a	-1	-1
121 ^a	-1	-1	166 ^a	-1	-1
122 ^a	-1	-1	167 ^a	-1	-1
123 ^a	-1	-1	168 ^a	-1	-1
124 ^a	-1	-1	169 ^a	-1	-1
125 ^a	-1	-1	170 ^a	0	0
126 ^a	-1	-1	171 ^a	0	0
127 ^a	-1	-1	172 ^a	0	0
128 ^a	0	0	173 ^a	1	1
129 ^a	0	0	174 ^a	1	1
130 ^a	0	0	175 ^a	1	1
131 ^a	0	0	176 ^a	1	1
132 ^a	1	1	177 ^a	1	1
133 ^a	1	1	178 ^a	1	1
134 ^a	1	1	179 ^a	1	1



O Impacto das Regras do Programa Bolsa Família Sobre a Fecundidade das Beneficiárias*

Luis Antonio Winck Cechin,[†] André Carraro,[‡] Felipe Garcia Ribeiro,[§]
Rodrigo Nobre Fernandez[¶]

Sumário: 1. Introdução; 2. Descrição e evolução do Programa Bolsa Família; 3. Modelo teórico de demanda por filhos; 4. Estratégia empírica; 5. Dados, população-alvo e limitações; 6. Análise dos resultados; 7. Considerações finais; Apêndice A. Prova das proposições do modelo teórico; Apêndice B. Tabelas.

Palavras-chave: Programa Bolsa Família, Fecundidade, Propensity Score Matching.

Códigos JEL: J13, J18, I38.

Este trabalho investiga um possível incentivo do Programa Bolsa Família ao aumento da fecundidade de suas beneficiárias em decorrência de suas regras, dado que a quantidade de recursos transferidos depende do número de filhos da família. O diferencial deste estudo reside na análise desse impacto em um maior período de exposição das beneficiárias aos efeitos do PBF. Aplica-se o algoritmo de seleção de covariadas proposto por Imbens (2014) e o método de *Propensity Score Matching*. Os resultados apontaram que o PBF gera pequeno incentivo à geração do segundo filho, sendo que as regiões Centro-Oeste e Nordeste apresentaram os maiores valores de impacto.

This paper investigates a possible incentive from the Bolsa Família Program to increase the fertility of its beneficiaries due to its rules, according to which the amount of funds transferred depends on the number of family children. The differential of this study lies in the analysis of this impact over a longer period of exposure of the beneficiaries to the effects of the Bolsa Família Program. Applies the covariates selection algorithm proposed by Imbens (2014) and the method of Propensity Score Matching. The results showed that the Bolsa Família Program generates little incentive to the generation of the second child, being the Midwest and Northeast regions had the highest impact values.

*Agradecemos ao CNPQ pelo suporte financeiro necessário ao desenvolvimento desta pesquisa. Agradecemos especialmente ao Professor André Portela Fernandes de Souza (FGV-EESP) pelas discussões e sugestões.

[†]Programa de Pós-Graduação em Organizações e Mercados, PPGOM/UFPEL. E-mail: lwn_cechin@hotmail.com

[‡]Departamento de Economia, Universidade Federal de Pelotas (UFPEL). E-mail: andre.carraro@gmail.com

[§]Departamento de Economia, Universidade Federal de Pelotas (UFPEL). E-mail: felipe.garcia.rs@gmail.com

[¶]Departamento de Economia, Universidade Federal de Pelotas (UFPEL). E-mail: rodrigo@rodrigofernandez.com.br



1. INTRODUÇÃO

O Programa Bolsa Família (PBF) é o maior Programa de Transferência Condicionada de Renda do mundo, beneficiando aproximadamente 13,4 milhões de famílias brasileiras em 2015.¹ Em 2013, o PBF completou 10 anos de vigência, tendo como objetivo proporcionar o alívio imediato da pobreza, auxílio à educação, saúde e assistência social e superação da situação de vulnerabilidade social das famílias pobres e extremamente pobres.

As regras do Programa determinam que a quantidade de recursos transferidos depende do número de filhos da família, ou seja, transfere mais renda para as famílias que têm mais filhos. Dessa forma, o Programa pode gerar um incentivo ao aumento da fecundidade de suas beneficiárias, tendo como consequência a multiplicação da população pobre. De acordo com Rocha (2009), a preocupação é que esses Programas possam gerar uma espécie de armadilha da pobreza.

O presente trabalho visa investigar se a taxa de fecundidade das beneficiárias é afetada pelas regras do Programa Bolsa Família. Este estudo contribui com a literatura econômica apresentando uma análise do efeito causal do PBF sobre a fecundidade, observando uma população que ficou seis anos e nove meses exposta aos efeitos do Programa. Utilizam-se os dados do Censo Brasileiro de 2010 e o método de *Propensity Score Matching* (por meio da metodologia do vizinho mais próximo, com suporte comum), combinado com as técnicas de *matching* propostas por Imbens (2014).

Em razão do seu desenho, o PBF pode incentivar a fecundidade das beneficiárias tanto pela ampliação da renda familiar quanto pela redução do custo de se ter um filho adicional. A decisão de gerar mais um filho dependerá da relação entre o custo com o filho adicional e a receita adicional oriunda do PBF. Dessa forma, se a receita marginal for maior que o custo marginal, a beneficiária poderá ser incentivada a gerar outro filho.

Apesar de o Programa Bolsa Família ter um desenho pró-natalista, a maioria dos estudos realizados até o momento aponta que o PBF não incentiva a fertilidade dos pobres. Rocha (2009) utilizou os microdados da PNAD, por meio dos métodos *triple-diff* (com as PNADs de 1995 a 2007), *Probit* (com a PNAD 2006) e *Propensity Score Matching* (com a PNAD 2006). Pelos três métodos, foi obtido resultado segundo o qual o PBF não tem efeito significativo na taxa de fecundidade das beneficiárias. Signorini & Queiroz (2011), utilizando as PNADs de 2004 e 2006, avaliaram o efeito médio do PBF na fertilidade das beneficiárias por meio do método *Propensity Score Matching*. Para ambos os anos, os autores não encontraram impactos significativos do recebimento do PBF na fertilidade das beneficiárias do programa.

Nessa mesma direção, Simões & Soares (2012) utilizaram a PNDS (Pesquisa Nacional de Demografia e Saúde) de 2006 e usando duas metodologias, dois-estágios estilo Heckman e Método dos Momentos Generalizados (GMM), verificaram que o Programa Bolsa Família não gerou incentivo à fecundidade. No método de GMM, os autores utilizaram como instrumento a variável de vizinhança, ou seja, o número de filhos que as vizinhas próximas possuíam.

No sentido oposto dos estudos anteriores, Berbel (2011), utilizando as PNADs de 1995, 1997, 1999 e de 2001 a 2007, usando o método de *Dif-in-Dif*, testou a hipótese de que a gravidez juvenil (mulheres entre 16 e 19 anos) estaria sendo incentivada por ser percebida como uma potencial porta de entrada para programa pela nova mãe. O autor concluiu que o PBF teve efeito positivo sobre a taxa de gravidez juvenil, gerando um efeito médio de 2,5 pontos percentuais.

Contudo, como o PBF teve início em outubro de 2003, esses estudos captaram apenas o efeito inicial de exposição ao Programa, pois utilizaram como fonte de dados basicamente as PNADs (Pesquisa Nacional de Amostra de Domicílios) de anos anteriores a 2003, do ano de 2004 e do ano de 2006, ou, ainda, a PNDS (Pesquisa Nacional de Demografia e Saúde) de 2006.

Dessa forma, esse curto período de análise pode não ser suficiente para avaliar o impacto do PBF sobre a variável de interesse, uma vez que os programas sociais podem ter impactos praticamente instantâneos sobre o mercado de trabalho, taxas de matrículas escolares, redução da desigualdade, etc.,

¹https://www.beneficiosociais.caixa.gov.br/consulta/beneficio/04.01.00-00_00.asp

mas não sobre a fecundidade. É razoável supor que os efeitos do Programa em relação à fecundidade não tenham impacto imediato, pois é possível que haja uma gradual compreensão das regras e dos benefícios do Programa por parte da população. Assim, evidencia-se a necessidade de um maior período de exposição das beneficiárias aos efeitos do PBF, de modo que seja possível avaliar se as decisões de fecundidade mudaram em função dos benefícios financeiros transferidos.

Nesse contexto, além de analisar um maior período de exposição aos efeitos do Programa (esse tipo de inferência já havia sido recomendada por Simões e Soares, 2012), este estudo se diferencia dos demais pela construção de um contrafactual mais adequado ao problema de interesse ao utilizar grupos de controle e tratamento situados dentro e fora da zona de incentivo² à fecundidade criada pelo Programa. Dessa forma, calcula-se a diferença das probabilidades das mulheres terem tido filho nascido vivo nos últimos doze meses, dado que as mulheres possuíam zero, um, dois, três ou quatro filhos em julho de 2009.

Diferentemente da estratégia de comparar a amostra de mulheres que possuíam dois filhos com a amostra de mulheres que possuíam três filhos (adotada por Rocha, 2009), ou de comparar a amostra de mulheres beneficiárias que possuíam menos de três filhos com a amostra das mulheres não beneficiárias, mas elegíveis ao PBF, que também possuíam menos de três filhos (adotada por Simões & Soares, 2012), este estudo compara a amostra de mulheres beneficiárias que possuíam zero, um, dois, três ou quatro filhos com a amostra de mulheres não beneficiárias, porém elegíveis ao PBF, que possuíam zero, um, dois, três ou quatro filhos.

Imbens (2014) recomenda a utilização de métodos *matching*, para se estimar inferência causal com robustez. Por meio da utilização dos dados de Lalonde (1986), Imbens aplicou suas técnicas em dados experimentais e não experimentais. Dessa forma, embora a maior crítica dos métodos *matching* resida no possível viés oriundo das variáveis não observáveis, os valores obtidos pelo autor com os dados não experimentais ficaram muito próximos dos valores obtidos com os dados experimentais, indicando a potência de suas técnicas em obter resultados robustos.

Neste estudo se adota a construção de um estimador de *propensity score* mais flexível, que utiliza variáveis lineares e de segunda ordem, as quais são escolhidas não por arbítrio do pesquisador, mas através de regressões logísticas e de testes de razão de máxima verossimilhança. Além disso, utilizando-se das técnicas de *matching* propostas por Imbens (2014), com o intuito de se obter resultados robustos, mesmo com dados não experimentais, foram realizados cortes na amostra para garantir melhor equilíbrio nas covariadas. Foram utilizados os dados do Censo Brasileiro de 2010, tendo sido adotada uma estratégia de decomposição da variável V0657,³ que possibilitou a identificação e separação dos beneficiários do PBF e do PETI.⁴

Os resultados obtidos indicam a existência de um efeito causal entre o Programa Bolsa Família e a taxa de fecundidade das beneficiárias, embora pouco expressivo. Para o grupo das mulheres que tinham apenas um filho em 2009, o PBF gerou um pequeno incentivo à fecundidade (1,31 pontos percentuais). Entre os “extremamente pobres”, o valor do efeito em nível nacional também foi baixo, pois uma beneficiária do PBF tem probabilidade 2,24 pontos percentuais maior de gerar o segundo filho em comparação com o seu *match* do grupo de controle. Além disso, as regiões Centro-Oeste e Nordeste apresentaram os maiores valores válidos desse impacto.

Por fim, este trabalho está dividido em seis seções além desta introdução. Na seção 2 será realizada a apresentação do PBF. Na seção 3 será apresentado um modelo teórico de demanda por filhos. A seção 4

²Em 2010, as mulheres que tinham menos de três filhos de até 15 anos pertenciam à zona de incentivo do Programa, pois se gerassem mais um filho, receberiam uma compensação financeira, e as mulheres que possuíam três filhos ou mais estavam fora da zona de incentivo, pois naquela época não havia recursos financeiros destinados ao quarto ou quinto filho.

³No Censo de 2010, para a variável V0657 foi atribuída a seguinte pergunta: “Em julho de 2010, tinha rendimento mensal habitual de Programa Social Bolsa-Família ou Programa de Erradicação do Trabalho Infantil (PETI)?”

⁴O Programa de Erradicação do Trabalho Infantil (PETI) visa erradicar o trabalho infantil (crianças de 7 a 15 anos) nas atividades perigosas, insalubres ou degradantes, nas zonas urbana e rural. Como condicionalidade, as crianças menores de 16 anos não podem trabalhar e devem participar da jornada escolar ampliada (dois turnos).



discutirá a estratégia empírica utilizada para atingir os objetivos propostos. A seção 5 versará sobre dados, população-alvo e limitações deste estudo. Na seção 6 será apresentada a avaliação empírica e os resultados encontrados. Após serão efetuadas as considerações finais.

2. DESCRIÇÃO E EVOLUÇÃO DO PROGRAMA BOLSA FAMÍLIA

O Programa Bolsa Família foi criado pela Medida Provisória nº 132, de 20 de outubro de 2003, a qual foi convertida na Lei 10.836, de 9 de janeiro de 2004. O PBF é responsável pela transferência direta de renda, beneficiando famílias que vivem nas situações de pobreza e de extrema pobreza. Para participar do PBF, as famílias interessadas devem possuir renda mensal per capita compatível com as regras do Programa e estarem devidamente cadastradas no Cadastro Único para Programas Sociais (CadÚnico).⁵

Ao longo da existência do PBF, a quantidade de benefícios transferidos às famílias e seus valores tiveram algumas mudanças e reajustes.⁶ Em 2010, período de análise deste estudo, uma família da categoria “extremamente pobre” (com renda familiar per capita de até R\$ 70,00) podia receber do PBF o Benefício Básico (que correspondia a R\$ 68,00), o Benefício Variável (que correspondia a R\$ 22,00 vezes o número de filhos de até 15 anos, até o limite de três filhos) e o Benefício Variável Vinculado ao Adolescente (BVJ, que correspondia a R\$ 33,00 vezes o número de filhos de 16 a 17 anos, até o limite de dois filhos); ou seja, uma família “extremamente pobre” poderia receber do PBF valores que variavam de R\$ 68,00 a R\$ 200,00.

Já uma família da categoria “pobre” (com renda familiar per capita de até R\$ 70,01 a R\$ 140,00) somente poderia participar do PBF se tivesse filhos com idade de até 17 anos, podendo receber o Benefício Variável (que correspondia a R\$ 22,00 vezes o número de filhos de até 15 anos, até o limite de três filhos) e o BVJ (que correspondia a R\$ 33,00 vezes o número de filhos de 16 a 17 anos, até o limite de dois filhos); ou seja, uma família “pobre” poderia receber do PBF valores que variavam de R\$ 22,00 a R\$ 132,00.

Atualmente, para participar do PBF, as famílias interessadas devem possuir renda mensal per capita de até R\$ 154,00. Entretanto, as famílias pobres (com renda familiar per capita entre R\$ 77,01 e R\$ 154,00) somente poderão participar do Programa se possuírem filhos (crianças ou adolescentes) de até 17 anos. Para as famílias extremamente pobres, a partir de março de 2013 (por força da Lei nº 12.817/2013), foi implantado o Benefício para Superação da Extrema Pobreza (BSP), que é calculado caso a caso, a fim de garantir que nenhuma família beneficiária tenha renda per capita inferior a R\$ 77,00.

Quando a família estiver inserida no Programa, sua permanência fica vinculada às condicionalidades de educação e saúde. No âmbito da educação, deve ser cumprida a frequência escolar mensal de 85% para crianças e de 75% para adolescentes, em estabelecimento de ensino regular. No âmbito da saúde, as condicionalidades são: o acompanhamento do desenvolvimento infantil e o acompanhamento das mulheres com idades entre 14 e 44 anos, bem como, no caso de gestantes e lactantes, a realização do pré-natal e o acompanhamento da saúde da mãe e do bebê.

O número de famílias beneficiárias do Bolsa Família cresceu significativamente ao longo dos seus dez anos de existência. De acordo com os dados da Controladoria Geral da União (CGU),⁷ em 2003 o PBF assistia a 3,6 milhões de famílias e a partir de 2012 esse número praticamente se estabilizou próximo dos 13,9 milhões. Por fim, além do crescente número de beneficiários, as principais mudanças no Programa, foram:

⁵Regulamentado pelo Decreto nº 6.135/07, o Cadastro Único identifica e distingue as famílias com renda mensal per capita de até meio salário mínimo ou de três salários mínimos no total. Dessa forma, o Cadastro Único traz várias informações acerca do núcleo familiar, possibilitando, portanto, a observação da realidade socioeconômica dessas famílias.

⁶As principais mudanças e reajustes podem ser verificadas em Souza, Oliveira, Duarte, Pereira & Gadelha, 2013, p.7.

⁷<http://www.portalttransparencia.gov.br>

- (i) Em 2005, a integração do PETI com o Programa Bolsa Família,⁸ a qual não objetivou a extinção do PETI, mas focou na sinergia dos dois programas para combater a pobreza e erradicar o trabalho infantil.
- (ii) Em 2008, a implantação do Benefício Variável Vinculado ao Adolescente (BVJ).
- (iii) Em 2011, a Medida Provisória nº 535, que aumentou o número de benefícios por filho, passando de três BV's para cinco BV's.
- (iv) Em 2013, a implantação do Benefício para Superação da Extrema Pobreza (BSP), criado para garantir que nenhum beneficiário do PBF continue em situação de extrema pobreza.

3. MODELO TEÓRICO DE DEMANDA POR FILHOS

Tendo como base o modelo de demanda por filhos de Becker (1960), foi adicionada a remuneração proveniente do Programa Bolsa Família na restrição orçamentária das famílias, bem como foi considerado que os pressupostos básicos na decisão do consumidor⁹ são válidos para a escolha entre ter filhos e consumir outros bens.

A decisão da família de gerar mais um filho depende intrinsecamente da relação de custo *versus* benefício que esta criança trará. Em termos formais, a função de utilidade da família representativa, definida $U: \mathbb{R}^2 \rightarrow \mathbb{R}$, relaciona a escolha entre o número de filhos n e o consumo de outros bens Z .¹⁰ Assume-se que a função de utilidade é quase-côncava, isto é, o incremento de n e Z aumenta a satisfação da família representativa a taxas decrescentes.

Para que esse agente econômico receba uma renda adicional proveniente do PBF é necessário que a renda per capita da unidade familiar seja inferior ou igual a R\$ 154,00. Primeiramente, define-se $Y: \mathbb{R}^3 \rightarrow \mathbb{R}$, que é a função que determina a renda familiar per capita e possui a seguinte forma:

$$Y = f(n, c, R) = \frac{R}{n + c},$$

onde R é a renda total da família, n é o número de filhos de até 17 anos e c é o número total de adultos.

O próximo passo é a construção da função que determina o benefício a ser recebido do programa. Para simplificar a análise, define-se $BF: \mathbb{R}^4 \rightarrow [bf_{\min}, bf_{\max}]$ como uma função que depende de j , k , Y e θ , ou seja, $BF = f(j, k, Y, \theta)$. O parâmetro j representa o número total de filhos até 15 anos, sendo que o valor do PBF destinado à j é definido como $j: \mathbb{R}^2 \rightarrow [j_{\min}, j_{\max}]$, que tem a seguinte forma $j = f(n, \theta) = j(n, \theta) = (1 - \theta)n\bar{j}$; ou seja, o PBF transfere o valor de R\$ 35,00 (\bar{j}) para cada filho de até 15 anos, porém cada família pode receber no máximo cinco benefícios.

Em seguida, tem-se que k é o total dos filhos que possuem idade de 16 ou 17 anos; desse modo, o valor do PBF destinado à k é definido como $k: \mathbb{R}^2 \rightarrow [k_{\min}, k_{\max}]$, sendo que $k = f(n, \theta) = k(n, \theta) = n\theta\bar{k}$; ou seja, o PBF transfere o valor de R\$ 42,00 (\bar{k}) para cada filho de 16 ou 17 anos, porém cada família pode receber no máximo dois benefícios. Note que $\theta \in [0, 1]$, em outras palavras, este parâmetro representa o percentual de filhos mais velhos. Adicionalmente supõe-se que $n \in [n_{\min}, n_{\max}]$, já que não há benefício adicional proveniente do PBF se o número total de filhos for maior que o limite superior (sete filhos). Em resumo, a função de benefício $BF = f(j, k, Y, \theta)$ pode ser definida da seguinte forma:

$$BF(j, k, Y, \theta) = b + j + k, \quad \forall n \in [n_{\min}, n_{\max}],$$

⁸ A integração do Programa Bolsa Família com o PETI foi regulamentada pela Portaria nº 666, de 28 de dezembro de 2005. http://www.mpggo.mp.br/portal/arquivos/2013/05/14/11_18_16_239_PETI_Portaria_n%C2%BA_666_de_28_de_dezembro_de_2005.pdf

⁹ i) A família sempre escolhe uma cesta de consumo sobre a reta orçamentária; ii) A família sempre consome toda a renda disponível. As preferências são completas, monótonas, transitivas, convexas e contínuas.

¹⁰ Z é um vetor de outros bens, formalmente definido como $Z = \{x \in \mathbb{R} \mid x > 0\}$. Para simplificar a análise, normalizou-se para um único bem.



sendo que

$$b = \begin{cases} 0, & \text{se } Y > 77 \\ 77, & \text{se } Y \leq 77 \end{cases}.$$

Além disso, a função de benefício do Bolsa Família será zero para o caso de o número de filhos ser igual a zero e a renda familiar per capita ser superior a R\$ 77,00, ou seja, uma família da categoria “pobre” somente será beneficiária do PBF se possuir ao menos um filho.

Uma hipótese existente neste modelo é a da mudança de categoria (de pobre para extremamente pobre), que ocorre em virtude do nascimento de novos filhos. Suponha que uma família esteja classificada na categoria “pobre” (renda per capita de R\$ 77,01 a R\$ 154,00), mas esteja próxima da faixa da extrema pobreza (renda per capita de até R\$ 77,00), assim, quando for gerado outro filho, a renda per capita familiar irá reduzir e a categoria da família será alterada para “extremamente pobre”. Essa mudança lhe garante um recurso financeiro maior, pois além de receber o benefício adicional por filho, também auferirá o benefício básico de R\$ 77,00. Essa condição cria um incentivo para que as famílias beneficiárias possam gerar novos filhos com o intuito de ampliar a renda familiar.

Outro ponto importante é a definição do custo de se ter um novo filho ser decrescente em n , ou seja, a estrutura de custos é marginalmente decrescente, além de existir efeitos de aprendizagem a cada nova criança gerada. Isso decorre em razão do *background* familiar, que leva em consideração as experiências passadas e a estrutura física remanescente e aproveitável dos filhos anteriores (móveis, roupas, etc.). No entanto, ao ficarem mais velhos os custos aumentam. Isto é, quando θ aumenta a função, o preço cresce. Desse modo, o preço do nascimento de uma criança é uma função $p: \mathbb{R}^2 \rightarrow \mathbb{R}_+$, que tem a forma $p = f(n, \theta)$.

A partir desses pressupostos, pode-se construir a restrição orçamentária da família representativa:

$$p(n, \theta)n + \pi Z = R + BF(j, k, Y, \theta),$$

onde π é o preço do bem Z e R é uma renda exógena gerada pelo trabalho ou outras fontes de renda da família. A partir daqui pode-se solucionar o problema de maximização da utilidade da família chegando ao seguinte resultado:

Proposição 1. *As demandas ótimas por n e Z são¹¹*

$$n = \frac{1}{p_n} \left(\frac{\pi U_n}{U_Z} + \phi \right) \quad e \quad Z = \frac{1}{\pi} \left[R + BF - \frac{p}{p_n} \left(\frac{\pi U_n}{U_Z} + \phi \right) \right].$$

Tem-se que

$$\phi = BF_j (1 - \theta) \bar{j} + BF_K \theta \bar{k} - p(n, \theta) - BF_K \frac{R}{(n + c)^2} \quad e \quad \frac{U_n}{U_Z} > \phi.$$

Nota-se que a demanda por outros bens depende negativamente da demanda por filhos. Em resumo, um filho adicional elimina, para os pais, a ideia de troca para o consumo de outros bens. Pode-se observar essa relação olhando para a taxa marginal de substituição entre o consumo de “filhos” *versus* outros bens:

$$\frac{U_n}{U_Z} = \frac{np_n - \phi}{\pi}.$$

Proposição 2. *A taxa marginal de substituição aumenta conforme o número de filhos cresce.*

¹¹Para pequenas variações de n . Maiores detalhes são apresentados no Apêndice A.

Tem-se que quando o número de crianças aumenta, torna-se mais caro trocar um filho pelo consumo de outros bens. Na cesta de consumo ótima, o valor da TMS é igual à razão dos preços relativos.

Proposição 3. *O benefício marginal oriundo do nascimento do primeiro filho é estritamente positivo para as famílias dos grupos de renda “pobres” ou “extremamente pobres”.*

Proposição 4. *A medida que o número de filhos aumenta, o benefício marginal recebido pelo PBF tende a zero.*

A Proposição 3 indica que, para as famílias não beneficiárias cuja renda se enquadre na categoria “pobre”, o nascimento do primeiro filho torna a família elegível ao PBF, pois uma família dessa categoria somente será beneficiária do PBF se possuir ao menos um filho. Certamente esse resultado vale para o nascimento da primeira criança, mas não necessariamente para os próximos filhos.

Finalmente, a Proposição 4 indica que se o número de filhos tende ao máximo, isto é, ao limite superior que proporcionará o recebimento de um benefício do programa; na margem, o efeito desse benefício é inferior ao efeito nos custos e ϕ pode se tornar negativo, embora muito pequeno. Em outras palavras, o *trade-off* da família aumenta entre ter um novo filho e consumir um conjunto de outros bens, definido por Z . No Apêndice A, encontram-se as provas de todas as proposições. Por fim, a próxima seção trata da estratégia empírica adotada neste estudo.

4. ESTRATÉGIA EMPÍRICA

É relevante destacar que quando se estudam os efeitos causais de uma política pública, surge um problema: só se consegue observar o factual, ou seja, aqueles indivíduos que foram selecionados ao Programa e, dessa forma, receberam o “tratamento” da política. No entanto, para que se possa estabelecer o efeito causal de determinada política pública, é necessário encontrar um grupo de controle que seja o mais próximo possível do contrafactual (como estaria o indivíduo “tratado” caso ele não tivesse recebido o “tratamento”).

O principal problema da avaliação de impacto de políticas públicas reside em encontrar um grupo de controle que represente adequadamente o contrafactual do grupo tratado. Embora o método experimental seja o ideal para se estimar o efeito causal com robustez, ele raramente é utilizado para selecionar famílias (ou indivíduos) em políticas sociais de transferência de renda. No caso do Programa Bolsa Família, os beneficiários não foram escolhidos por um processo de seleção aleatória. Assim, se incorre no problema de autosseleção¹² dos indivíduos em participar do PBF.

Dessa maneira, a busca de uma relação causal entre a participação no Programa e as mudanças nas variáveis de interesse carece da análise de uma população não beneficiária do PBF, porém, comparável com a dos beneficiários. Para isso, é necessário encontrar famílias não beneficiárias que sejam semelhantes aos participantes dos programas, ou seja, que possuam as mesmas características das famílias beneficiárias (equivalência de renda, UF, local de residência, cor, escolaridade, etc.), e que a diferença de comportamento entre elas seja atribuída tão somente no que se refere ao recebimento do PBF.

4.1. Propensity Score Matching

O método de *Propensity Score Matching* consiste em buscar um “grupo de não tratados” semelhante ao “grupo dos tratados” baseando-se num vasto conjunto de características observáveis dos indivíduos. No entanto, ao invés de realizar o pareamento com todos os indivíduos com base em toda a dimensão da matriz X , Rosenbaum & Rubin (1983) sugeriram que os indivíduos fossem pareados utilizando uma função de X , a qual calcula a probabilidade do indivíduo receber o tratamento dado o conjunto de características contidas em X . Dessa forma, o *propensity score* faz com que se reduza a dimensionalidade das

¹²Sobre o problema de autosseleção, ver Heckman (1979).



variáveis da matriz X . Por fim, através da matriz X e do escore de propensão, quanto mais semelhantes forem os dois grupos (tratamento e controle) em decorrência das características observáveis, mais a variável dependente passa a independender da participação ou não no tratamento, o que suaviza o viés de seleção.

Para que as estimações sejam viáveis, o modelo de pareamento utiliza duas hipóteses: i) *unconfoundedness* (também conhecida como seleção em observáveis, exogeneidade, ignorabilidade, ou simplesmente independência condicional); e ii) sobreposição. Essas hipóteses postulam que ao se comparar um indivíduo do grupo de tratamento com um indivíduo do grupo de controle, cujas características observáveis de ambos são iguais, o único fator que diferencia os resultados observados desses indivíduos seria o recebimento ou não do tratamento. De acordo com Imbens (2014), cada unidade na amostra é observada para receber ou não um tratamento binário, com o indicador de tratamento designado por W_i . Assim, se a unidade i recebe o tratamento então $W_i = 1$, caso contrário $W_i = 0$.

O estimador de *propensity score*, $\hat{e}(x)$, é baseado no modelo de regressão logística, estimado por máxima verossimilhança. Dado a escolha para a função $h(x)$ e a estimação do parâmetro desconhecido γ por máxima verossimilhança, o estimador de *propensity score* será

$$\hat{e}(x | W, X) = \frac{\exp\left(h(x)' \hat{\gamma}_{ml}(W, X)\right)}{1 + \exp\left(h(x)' \hat{\gamma}_{ml}(W, X)\right)} \quad (1)$$

A escolha das variáveis para compor a função $h(x)$ foi feita com o auxílio do algoritmo de seleção de covariadas proposto por Imbens (2014).¹³ Esse algoritmo proporciona uma especificação mais flexível para o *Propensity Score*, pois seleciona covariadas lineares e de segunda ordem (variáveis quadráticas e variáveis de interação entre as covariadas lineares) para compor o escore de propensão. Além disso, as variáveis são escolhidas não por arbítrio do pesquisador, mas com base em testes de razão de máxima verossimilhança.

O escore de propensão é composto por uma variável linear pré-tratamento (número de filhos em julho de 2003, que representa um período anterior à implantação do Programa Bolsa Família), por termos lineares (idade, renda familiar per capita *ex-ante*, densidade de pessoas por cômodos, *dummies* de cor, urbano/rural, se tem cônjuge, se é chefe do domicílio, *dummies* de escolaridade — baixa instrução, ensino fundamental completo, ensino médio completo —, se possui energia elétrica, se tem aposentado ou pensionista no domicílio, e *dummies* de estado — unidade da federação em que reside a mulher), bem como por termos de segunda ordem (interações entre as variáveis anteriormente relacionadas).

Neste trabalho, se utiliza a renda familiar per capita *ex-ante*,¹⁴ pois, além de ser a variável de rendimentos adequada para se estimar o *Propensity Score*, é a variável que possibilita a separação das mulheres nas duas categorias de beneficiárias do PBF (“pobres” e “extremamente pobres”). Essa separação é importante para se captar o efeito do PBF sobre a fecundidade em cada categoria do Programa.

Sendo assim, sob as hipóteses de independência condicional e sobreposição, o estimador de *propensity score matching* para o efeito médio do tratamento sobre os tratados (ATT), τ_t , pode ser definido como

$$\hat{\tau}_{t,N} = \frac{1}{N_t} \sum_{i=1}^N W_i \left(Y_i - \frac{1}{M} \sum_{j \in J_M(i)} Y_j \right), \quad (2)$$

em que M é o número de *matches* por unidade e $J_M(i)$ é o conjunto de *matches* por unidade i e $N_t = \sum_{i=1}^N W_i$ é o número unidades tratadas na amostra.

¹³Os detalhes para a construção do algoritmo de seleção de covariadas podem ser obtidos em Imbens (2014).

¹⁴A renda familiar per capita *ex-ante* desconta do rendimento total familiar os valores oriundos do PBF. O método de construção da variável “renda familiar per capita *ex-ante*” para os dados do Censo Brasileiro de 2010 pode ser solicitado por e-mail, juntamente com os códigos para o software Stata.

Isso posto, para realizar o pareamento e estimar o efeito médio do tratamento sobre os tratados de forma robusta, foi necessário realizar as três etapas indicadas por Imbens (2014): i) avaliar a hipótese de sobreposição; ii) avaliar a plausibilidade da hipótese de independência condicional; e iii) realizar a etapa de análise do *matching*.

Na primeira etapa, focando apenas no indicador de tratamento e na matriz de covariadas, (X, W) , foi necessário recortar o total da amostra, descartando algumas unidades com o intuito de melhorar a sobreposição na distribuição das covariáveis. Na segunda etapa, novamente focando apenas em (X, W) , foi verificada a plausibilidade da hipótese de independência condicional, através da estimação do efeito causal do tratamento em um “pseudorresultado”, ou seja, em uma variável conhecida por não ser afetada pelo tratamento, tendo em vista que o seu valor foi determinado antes da implantação do Bolsa Família. Por fim, na última etapa, utilizando a variável de interesse, Y , o estimador do efeito médio do tratamento sobre os tratados, $\hat{\tau}_{t,N}$, foi aplicado numa amostra cortada $\hat{\tau}_{t,N} = \tau(Y^T, W^T, X^T)$, para análise do *matching*.

É importante destacar que o efeito do tratamento estimado neste estudo utiliza erro padrão robusto,¹⁵ ou seja, leva em consideração o fato de que os escores de propensão são estimados em vez de verdadeiros ao calcular o erro padrão. Isso implica, em muitos casos, em uma diferença significativa na robustez da estimação.

Em todos os casos, a análise foi feita considerando apenas as mulheres com idade de 16 a 49 anos, ou seja, mulheres em idade reprodutiva¹⁶ e com idade suficiente¹⁷ para receber benefícios de Políticas de Transferência Condicionada de Renda, na condição de chefes de domicílio ou cônjuges do chefe, tendo em vista que são estas as responsáveis pelo recebimento do benefício e, em tese, poderiam ser incentivadas pelo Programa a gerar outro filho.

4.2. Análise de sensibilidade (limites de Rosenbaum)

Se a participação no PBF for endógena, ou seja, se variáveis não observáveis afetam tanto a participação no Programa quanto os resultados (no caso deste estudo, a fecundidade das beneficiárias), os estimadores do efeito médio do tratamento poderão não ser consistentes. Sendo assim, torna-se necessário avaliar o impacto potencial do viés de seleção, que decorre justamente das variáveis não observáveis.

Rosenbaum (2010) sugere avaliar o quão forte é a influência de uma variável não observada sobre a participação no Programa, a fim de prejudicar as conclusões do efeito causal do tratamento obtidas por meio dos métodos de pareamento. Nesse contexto, a análise de sensibilidade aplicada neste estudo é utilizada para testar a robustez dos resultados obtidos, considerando a presença de viés nos resultados (oriundo de uma covariável omitida).

A probabilidade de participação no PBF de um indivíduo i pode ser definida como

$$\pi_i = E[W_i | X_i = x] = \Pr(W_i = 1 | X_i = x) = F(\beta X_i + \gamma u_i). \quad (3)$$

Na ausência de viés de seleção, γ será igual a zero e a probabilidade de participação no PBF será determinada apenas pelas covariadas observáveis. Entretanto, se existir viés de seleção, dois indivíduos com as mesmas covariadas observáveis X terão diferentes probabilidades de receber o tratamento. Nessas condições, assumindo que $F(\cdot)$ tenha uma distribuição logística, a razão de probabilidades de dois

¹⁵Para tais estimativas foi utilizado o pacote *Teffects Psmatch* do software STATA 13. Esse pacote foi construído observando as orientações de Abadie & G.Imbens (2011); Abadie & Imbens (2006, 2008, 2012).

¹⁶Segundo informações do Ministério da saúde, a idade reprodutiva das mulheres brasileiras está compreendida entre o período de 10 a 49 anos (http://conselho.saude.gov.br/ultimas_noticias/2007/politica_mulher.pdf).

¹⁷O MDS estabelece a idade mínima de 16 anos para a inserção do responsável pela Unidade Familiar no CadÚnico (<http://www.mds.gov.br/falemds/perguntas-frequentes/bolsa-familia/cadastro-unico/gestor/cadunico-institucional>).



indivíduos pareados, i e j , que possuem as mesmas características observáveis ($X_i = X_j$), receberem tratamento é dada por

$$\frac{\frac{\pi_i}{1-\pi_i}}{\frac{\pi_j}{1-\pi_j}} = \frac{\pi_i(1-\pi_j)}{\pi_j(1-\pi_i)} = \frac{\exp(\beta X_j + \gamma u_j)}{\exp(\beta X_i + \gamma u_i)} = \exp[\gamma(u_i - u_j)]. \quad (4)$$

Se a razão de probabilidades for diferente de 1, fica configurada a presença de viés de seleção. Isso ocorre quando há diferenças nas variáveis não observáveis ($u_i \neq u_j$) e quando essas variáveis influenciam na probabilidade de participação ($\gamma \neq 0$). Em suma, a análise de sensibilidade avalia quanto que o efeito médio do tratamento é alterado pela mudança nos valores de γ e de $u_i - u_j$.

Rosenbaum (2002) sugere examinar os limites da razão de probabilidades de participação no tratamento. O autor mostra que a equação (4) implica nos seguintes limites:

$$\frac{1}{\Gamma} \leq \frac{\pi_i(1-\pi_j)}{\pi_j(1-\pi_i)} \leq \Gamma, \quad (5)$$

em que $\Gamma = e^\gamma$. Dessa forma, os indivíduos pareados possuem idêntica probabilidade de participação somente se $\Gamma = 1$. Entretanto, se $\Gamma > 1$, indivíduos aparentemente idênticos em termos de variáveis observáveis irão diferir em suas probabilidades de receber o tratamento. Por exemplo, se $\Gamma = 2$, a probabilidade de receber o tratamento será diferente por um fator de até 2. Por fim, os limites de Rosenbaum informam o quanto as mudanças nos valores de Γ alteram a inferência acerca do efeito do tratamento, estimado pelo método de *Propensity Score Matching*.

5. DADOS, POPULAÇÃO-ALVO E LIMITAÇÕES

Neste trabalho foram utilizados os dados do Censo 2010 do IBGE (Instituto Brasileiro de Geografia e Estatística), por se tratar da última fonte aberta de dados em nível domiciliar que aborda ambos os Programas: Bolsa Família e PETI. A possibilidade de identificação desses programas vem da variável V0657, cuja pergunta foi: “Em julho de 2010, tinha rendimento mensal habitual de Programa Social Bolsa Família ou Programa de Erradicação do Trabalho Infantil (PETI)?”

Dessa forma, para se identificar os beneficiários de cada programa, foi adotada uma estratégia de decomposição da variável V0657 em duas variáveis: “PBF” e “PETI”. É imprescindível saber que, em 2010, o Programa Bolsa Família estava presente em todas as cidades brasileiras e que o PETI estava presente em apenas 2.966 cidades. Logo, existiam municípios PETI e municípios não-PETI.

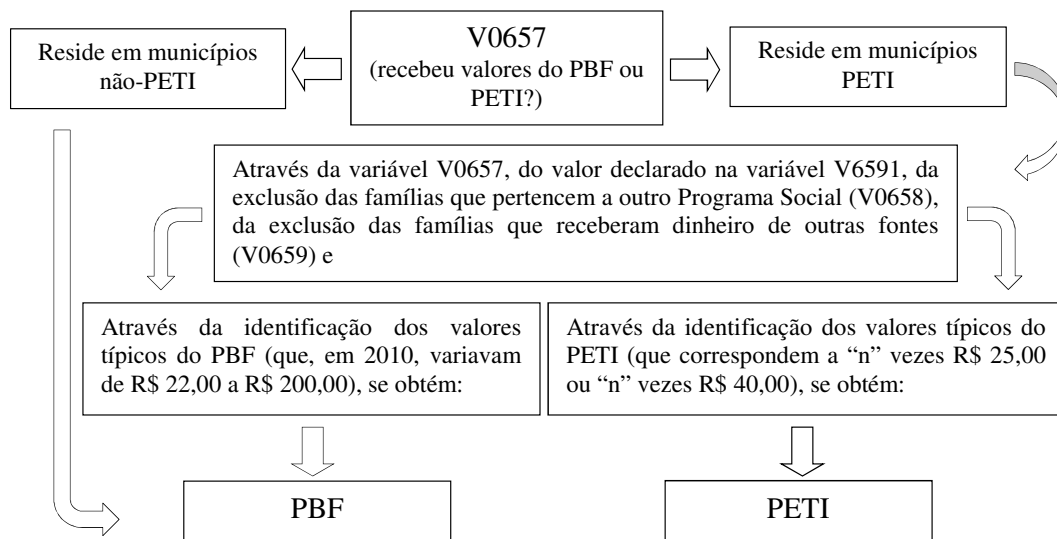
Sendo assim, essa decomposição só foi possível devido: i) às regras de ingresso aos Programas (as quais não permitem o recebimento de ambos os benefícios); ii) aos valores de transferência de cada programa (que eram diferentes);¹⁸ iii) à existência da variável V6591, cuja pergunta foi: “Em julho de 2010 qual foi o valor total deste(s) rendimento(s)?”; e iv) à identificação dos municípios PETI e dos municípios não-PETI em 2010.

Após a decomposição da variável V0657 do Censo de 2010 foram identificadas 614.208¹⁹ mulheres com idade de 16 a 49 anos (na qualidade de chefes do domicílio ou cônjuges do chefe) beneficiárias do Bolsa Família. Destas, somente 371.058 possuíam renda familiar per capita *ex-ante* de até R\$ 140,00, que representam 60,41% do total das beneficiárias identificadas. Porém, quando consideramos a renda per capita *ex-ante* de até R\$ 280,00, foram identificadas 534.073 mulheres, que representam 86,95% do total das beneficiárias identificadas na referida decomposição.

¹⁸Em 2010, existia apenas uma coincidência de valores em ambos os Programas: o valor de R\$ 200,00. Nesse caso, a separação/identificação dos beneficiários foi feita por meio do número de filhos de até 15 anos.

¹⁹Se forem considerados os pesos do CENSO 2010, o número de mulheres identificadas com idade de 16 a 49 anos, na qualidade de chefes do domicílio ou cônjuges do chefe, totaliza 4.067.455, sendo 2.357.299 com renda familiar per capita *ex-ante* de até R\$ 140,00 e 3.488.680 com renda familiar per capita *ex-ante* de até R\$ 280,00.

Figura 1. Decomposição da variável V0657.



5.1. Identificação da população-alvo e limitações

O impacto inicial do PBF sobre a fecundidade das beneficiárias pode ser observado através da PNAD 2006 (com período de 26 meses: de julho de 2004 a setembro de 2006) ou da PNDS 2006 (com período de 29 meses: de julho de 2004 a dezembro de 2006). No entanto, esse período pode ser muito curto para se captar o verdadeiro impacto do Programa sobre a fecundidade, seja por que o PBF aumentou significativamente o número de beneficiários em relação ao período inicial, seja porque esse tipo de incentivo não tem efeito instantâneo (ventila-se uma gradual compreensão, por parte da população beneficiária, das regras e dos possíveis benefícios do Programa).

Sendo assim, um maior período de exposição aos efeitos do Programa parece ser necessário para que se consiga captar o impacto do PBF sobre a fecundidade. Dessa forma, tal impacto pode ser avaliado através do CENSO 2010, que conta com uma população a qual foi exposta aos efeitos do PBF pelo período de seis anos e nove meses. No entanto, algumas limitações são impostas com os dados do CENSO 2010. A primeira limitação é referente à dificuldade de se identificar os beneficiários do Bolsa Família no CENSO, pois, mesmo tendo profundo conhecimento de ambos os Programas, foi possível identificar apenas 77,47% dos beneficiários do PBF ou PETI.

A segunda limitação reside no fato de que desconhecemos a data exata em que a beneficiária do PBF passou a ser “tratada”, ou seja, não é possível identificar, por meio do Censo de 2010, o mês/ano em que a família se tornou beneficiária do PBF. Entretanto, sabe-se que em 2010 a situação da família era “beneficiária” ou “não beneficiária”, bem como que o número de beneficiários teve apenas um pequeno acréscimo²⁰ entre 2009 e 2010. A terceira limitação é justamente em relação à fecundidade, pois não é possível identificar se a mulher teve filho nascido morto no período de referência de 12 meses anteriores a 31/07/2010.

A última limitação se refere à impossibilidade de se identificar no Censo 2010 a idade dos filhos que abandonaram o domicílio (seja pra constituir nova família, estudar/trabalhar em outro município/país ou, ainda, por qualquer outro motivo), bem como à impossibilidade de se identificar a idade dos filhos

²⁰Essa informação pode ser confirmada através do site <http://www.portaltransparencia.gov.br>.



que faleceram antes de 2009. Dessa forma, o presente estudo limitou-se a selecionar somente as famílias que “não reduziram”, ou seja, foram selecionadas apenas as famílias cuja quantidade de filhos tidos pela mulher (identificados através da variável V6800) foi igual à quantidade de filhos residentes e identificados no domicílio em 2010.

Apesar das limitações suprarreferidas, este trabalho possui a vantagem de observar o comportamento de fecundidade de uma população que ficou exposta aos efeitos do Programa por um período de seis anos e nove meses, sendo, possivelmente, tempo suficiente para que a população tivesse conhecimento do Programa, compreensão de suas regras e capacidade de avaliar seus benefícios. Além disso, com a utilização do algoritmo de seleção de covariadas e com o método sugerido por Imbens (2014), acredita-se ter construído um grupo de controle que representa adequadamente o contrafactual do grupo tratado.

Nesse contexto, a fim de evitar algum ruído nas estimativas, foram excluídas as mulheres cuja idade do filho mais novo era maior que 17 anos, em 2010. Essa medida é plausível porque para ser elegível ao PBF a mulher deve ter filhos de até 17 anos. Assim, ao se analisar a fecundidade das mulheres que tinham filhos de até 17 anos, verifica-se o comportamento reprodutivo da população elegível ao Programa.

Por fim, a população-alvo deste estudo são as mulheres com idade entre dezesseis e quarenta e nove anos, na qualidade de chefes do domicílio ou cônjuges do chefe, que se encontravam na zona de incentivo do Programa, ou seja, são as beneficiárias do Bolsa Família que tinham menos de três filhos em 2009 e poderiam ser incentivadas a gerar um filho a mais para ampliar a renda oriunda do Programa.

5.2. Detalhes das amostras e construção do contrafactual adequado

Para a construção do contrafactual adequado, foram selecionadas cinco amostras distintas, as quais foram separadas de acordo com o número de filhos das mulheres em 2009, cujos valores variam de zero a quatro filhos. Cada amostra foi novamente dividida em mais cinco amostras, desta vez sendo selecionadas pelo critério da renda familiar per capita *ex-ante*, o que possibilitou analisar o comportamento de fecundidade das mulheres com renda per capita *ex-ante* de até R\$ 70,00 (mulheres “extremamente pobres”), com renda per capita *ex-ante* de R\$ 70,01 até R\$ 140,00 (mulheres “pobres”), com renda per capita *ex-ante* de R\$ 0,00 até R\$ 140,00 (limite de elegibilidade ao Programa), com renda per capita *ex-ante* de R\$ 140,01 até R\$ 280,00, e, por fim, com renda per capita *ex-ante* de R\$ 0,00 até R\$ 280,00 (duas vezes o limite de elegibilidade ao Programa).

Em cada uma das vinte e cinco amostras, primeiramente, foi utilizado o algoritmo de seleção de covariadas,²¹ para que fossem selecionadas as variáveis lineares e de segunda ordem que melhor representassem cada amostra. Na sequência, foram estimados os escores de propensão, por máxima verossimilhança, e analisadas as estatísticas descritivas entre os grupos de tratamento e controle.

O equilíbrio das covariadas foi analisado através da diferença normalizada na média, ou seja,

$$\Delta_{ct} = \frac{\bar{X}_t - \bar{X}_c}{\sqrt{\frac{S_t^2 + S_c^2}{2}}}$$

De acordo com Imbens (2014), a diferença normalizada é mais útil que a estatística *t* para verificar a diferença das médias entre dois grupos, porque em grandes amostras a estatística-*t* pode ser grande, em valor absoluto, simplesmente porque a amostra é grande. De acordo com Imbens & Wooldridge (2009), quanto mais próximo de zero for o módulo da diferença normalizada, menor é a diferença entre a média dos grupos, sendo aceitável um valor de até 0,25.

De acordo com Imbens (2014), uma possibilidade de se obter maior equilíbrio nas covariadas é realizar a exclusão de observações com valores extremos de escore de propensão (próximos de zero ou um).²²

²¹Vide Imbens (2014) para detalhes.

²²Um aprofundamento desse método pode ser obtido em Crump, Hotz, Imbens & Mitnik (2008).

Imbens sugere deixar na amostra somente observações com valores de *pscore* maior que 0,1 e menor que 0,9. Então, para todas as amostras, foram excluídas as observações com valores de *pscore* menor que 0,1 e maiores que 0,9.

Como o objetivo deste trabalho é observar o efeito do tratamento sobre os tratados, além do corte nos escores de propensão, foi realizado o pareamento sem substituição (antes de se estimar o ATT), ou seja, cada unidade tratada foi pareada com uma única unidade de controle com similar valor de *pscore* (sendo tolerada uma diferença de no máximo 0,01). No entanto, o pareamento sem substituição só ocorreu nos casos em que o número de observações de controle era superior ao número de observações de tratados.

Nesse caso, após o pareamento restou um conjunto de mulheres para cada amostra, sendo metade do grupo de controle e metade do grupo de tratamento, onde cada unidade tratada tinha um par no grupo de controle com idêntico número de filhos em 07/2009, mesma renda per capita e similar valor de escore de propensão. Assim, sobre essas amostras cortadas, foi utilizado novamente o algoritmo de seleção de covariadas e reestimados os escores de propensão, por máxima verossimilhança. Posteriormente, foi estimado o efeito médio do tratamento em cada amostra.

Por fim, para as amostras em que, após o corte nos valores extremos de *pscore*, o número de observações de controle foi inferior ao de observações de tratados, foi realizado o pareamento com substituição, onde cada observação de controle pode ser pareada múltiplas vezes, e vice-versa. Essa medida foi utilizada para garantir maior representatividade entre os grupos.

6. ANÁLISE DOS RESULTADOS

Primeiramente, é importante ressaltar que não se pode falar em impacto causal do PBF sobre a fecundidade das beneficiárias considerando apenas o número total de filhos tidos pela mulher em 2010. A razão para isso reside no fato de que, por meio dos microdados do CENSO 2010, não se tem a informação da data em que cada mulher começou a participar do PBF e, portanto, começou a “receber o tratamento”. Sabemos apenas quais mulheres, em 2010, eram beneficiárias ou não do Bolsa Família.

Nesse contexto, para se estabelecer uma relação causal entre o PBF e fecundidade foi necessário verificar a fecundidade das mulheres nos doze meses anteriores ao Censo 2010. Desse modo, obtêm-se as duas informações necessárias para se estimar o efeito médio do tratamento: i) nascimento de filho(s) entre julho de 2009 e julho de 2010; e ii) identificação da situação das mulheres em relação ao Programa (se era beneficiária ou não).

Em 2010, o PBF pagava no máximo três BV's (um por filho). Sendo assim, após o terceiro filho, os nascimentos ocorreriam por outros motivos e não por incentivo do programa. Dessa forma, inicialmente foi estimado o impacto global do PBF sobre a fecundidade, ou seja, foi verificado o comportamento de fecundidade dentro da zona de incentivo (um ou dois filhos em julho de 2009) e fora da zona de incentivo (três filhos ou mais em julho de 2009).

Nesse primeiro momento, não foram consideradas as mulheres que não possuíam filhos em 2009, pois, com exceção das famílias extremamente pobres, para ser beneficiária do PBF a família tem que ter ao menos um filho. Além disso, apesar de utilizar uma metodologia diferente, esse recorte possibilita a comparação deste estudo com o trabalho de Simões & Soares (2012).

Os resultados iniciais apresentam efeitos negativos e/ou próximos de zero do PBF sobre a fecundidade, inclusive na suposta zona de incentivo criada pelo Programa. Esses resultados apontam para a ausência de efeito ou para um desincentivo do PBF à fecundidade. De acordo com a Tabela 1, o maior valor negativo do efeito do tratamento pertence às mulheres da categoria “pobre” (renda per capita de R\$ 70,01 a R\$ 140,00), que apresentam um ATT de $-1,05$ pontos percentuais, ou seja, uma beneficiária do PBF (com menos de três filhos) tem probabilidade 1,05 pontos percentuais menor de gerar outro filho, em comparação com o seu *match* do grupo de controle.

No entanto, em termos de probabilidade, esse valor é muito pequeno e não pode ser considerado como um desincentivo à fecundidade. Como os valores de ATT se encontram próximos de zero, a interpretação correta dos resultados da Tabela 1 é de ausência de efeito do PBF sobre a fecundidade. Dessa

**Tabela 1.** EFEITO DO TRATAMENTO SOBRE OS TRATADOS (BRASIL).

Existência de filho tido nascido vivo no período de referência de 12 meses anteriores a 31/07/2010 — a variável dependente assume valor igual a um se a mulher teve filho nesse período, e zero caso contrário.

Renda familiar per capita <i>ex-ante</i>	Mulheres com 1 ou 2 filhos em 07/2009 ($1 \leq n \leq 2$) coef/robust SE	Mulheres com 3 filhos ou mais em 07/2009 ($n \geq 3$) coef/robust SE
Até R\$ 70,00	0,0088*** (0,0009)	0,0011 (0,0008)
Nº de observações	461.458	451.805
De R\$ 70,01 a R\$ 140,00	-0,0105*** (0,0014)	-0,0092*** (0,0008)
Nº de observações	488.868	479.519 [†]
De R\$ 140,01 a R\$ 280,00	-0,0074*** (0,0006)	-0,0026*** (0,0006)
Nº de observações	849.191	395.839
De R\$ 0,00 a R\$ 140,00	-0,0001 (0,0008)	-0,0060*** (0,0005)
Nº de observações	829.078	976.688
De R\$ 0,00 a R\$ 280,00	-0,0080*** (0,0005)	0,0008 (0,0005)
Nº de observações	1.950.081	1.063.815

Notas: Erro padrão robusto entre parênteses. *p*-valor: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$. [†] indica a presença de desequilíbrio nas covariadas. As observações em **negrito** se referem às populações em que, após o corte dos valores extremos de *Propensity Score*, foi realizado o pareamento sem substituição, para, posteriormente, se estimar o efeito do tratamento. Nas demais, foram realizados apenas o corte dos valores extremos de *pscore*.

forma, esses resultados corroboram os trabalhos realizados anteriormente, os quais encontraram valores negativos ou ausência de efeito do PBF sobre a fecundidade (Rocha, 2009; Signorini & Queiroz, 2011; Simões & Soares, 2012).

Em busca de uma análise mais detalhada, foram estimadas vinte e cinco equações para verificar o impacto do Programa Bolsa Família sobre a fecundidade das beneficiárias. Dessa forma, calculou-se a nível nacional a diferença nas seguintes probabilidades:

- (i) Dado que a mulher não tinha nenhum filho em julho de 2009, qual é a probabilidade de ter tido o primeiro filho até julho de 2010?
- (ii) Dado que a mulher tinha um filho em julho de 2009, qual é a probabilidade de ter tido o segundo filho até julho de 2010?
- (iii) Dado que a mulher tinha dois filhos em julho de 2009, qual é a probabilidade de ter tido o terceiro filho até julho de 2010?
- (iv) Dado que a mulher tinha três filhos em julho de 2009, qual é a probabilidade de ter tido o quarto filho até julho de 2010?
- (v) Dado que a mulher tinha quatro filhos em julho de 2009, qual é a probabilidade de ter tido o quinto filho até julho de 2010?

Conforme relatado nas seções 4.1 e 5.2, para se estimar o efeito médio do tratamento sobre os tratados de forma robusta, foi necessário, inicialmente, descartar algumas unidades com o intuito de melhorar a sobreposição na distribuição das covariáveis. Através das figuras da Tabela B-2 do Apêndice A,

é possível verificar a hipótese de sobreposição entre os grupos de tratamento e controle. Observa-se a importância de se realizar os cortes nos escores de propensão para se obter uma amostra mais equilibrada, pois para as amostras em que foi possível fazer o pareamento sem substituição após o corte nos valores extremos de *pscore* (por exemplo, as figuras da zona de incentivo do PBF), se observa uma ótima sobreposição dos grupos, ou seja, a região da matriz *X*, que engloba as características observáveis dos indivíduos tratados também representa as características observáveis dos indivíduos que estão no grupo controle. Cabe ressaltar que todas as estimativas realizadas neste trabalho foram feitas sobre amostras reduzidas.

Posteriormente, foi verificada a plausibilidade da hipótese de independência condicional, sendo estimado o efeito causal do tratamento sobre um pseudorresultado. Para tanto, como variável dependente foi utilizada a variável pré-tratamento “número de filhos em 2003” e como variáveis independentes as mesmas covariadas escolhidas pelo algoritmo de seleção para se estimar o efeito do tratamento na variável de interesse. O único grupo de mulheres em que não foi possível verificar a viabilidade dessa hipótese foi o grupo das mulheres que não tinham filhos em julho de 2009, pois, obviamente, todas as observações tinham valor zero para 2003.

Por meio dos resultados da Tabela B-3 do Apêndice A, é possível constatar que a hipótese de independência condicional é plausível para poucas amostras (apenas para as amostras cujo coeficiente está em negrito), pois as estimativas ou tiveram valores próximos de zero ou não foram estatisticamente significantes. Por exemplo, para a amostra de mulheres que tinham dois filhos em 2009, com renda familiar per capita *ex-ante* de R\$ 70,01 até R\$ 140,00, o pseudorresultado foi de 0,56 pontos percentuais. Já para as mulheres com renda familiar per capita *ex-ante* de até R\$ 280,00, o pseudorresultado foi de 0,12 pontos percentuais, com significância de 1%.

A hipótese de independência condicional não se mostra plausível nas demais amostras da Tabela B-3, pois os valores de ATT indicam efeito do tratamento numa variável que foi determinada antes do Programa. Por exemplo, a amostra das mulheres com dois filhos e renda familiar per capita *ex-ante* de até R\$ 70,00 apresentou um ATT de 3,56 pontos percentuais (significante a 1%). No entanto, em amostras não experimentais, a ausência de evidências da plausibilidade da hipótese de independência condicional não pode ser enquadrada como problema, pois Imbens (2014) também não encontrou na utilização dos dados de Lalonde (1986).

Após a verificação das hipóteses de sobreposição e independência condicional, passou-se à análise da variável objeto do presente estudo. A Tabela 2 apresenta os resultados de todas as amostras. Observa-se que o maior valor do ATT foi 2,24 pontos percentuais, ou seja, uma beneficiária do PBF (da categoria “extremamente pobre”) tem probabilidade 2,24 pontos percentuais maior de gerar o segundo filho em comparação com o seu *match* do grupo de controle. Essa diferença de 2,24 pertence justamente à zona de incentivo do Programa, sendo que após o terceiro filho, quando a mulher deixa de ganhar um adicional para cada filho gerado, a diferença é praticamente zero.

Na Tabela 2, se observam valores negativos para algumas amostras, o que pode ser um indicativo de que o PBF, para esses casos específicos, gera um desincentivo à fecundidade. Já para a população extremamente pobre (com renda per capita de até R\$ 70,00), o efeito do tratamento é positivo para todas as amostras, demonstrando a fragilidade dessa população frente a uma pequena variação na renda (advinda do PBF).

A comparação destes resultados com os outros trabalhos presentes na literatura nacional (Rocha, 2009; Signorini & Queiroz, 2011; Simões & Soares, 2012) fica prejudicada em razão da estratégia de separação dos grupos de controle e tratamento, conforme discutido na introdução deste trabalho.

Com exceção das amostras das mulheres com quatro filhos e com renda per capita de até R\$ 140,00 e de até R\$ 280,00, todas as demais estavam devidamente equilibradas nas covariadas. Sendo assim, o acréscimo da fecundidade no limite do terceiro filho é pequeno em todos os casos, pois o maior valor de ATT foi 1,41 pontos percentuais, indicando que, embora os valores de ATT sejam positivos e estatisticamente significativos, os seus valores em nível nacional são baixos, apresentando baixa relevância econômica no que tange ao incentivo da fecundidade em relação às beneficiárias.

Tabela 2. EFEITO DO TRATAMENTO SOBRE OS TRATADOS (BRASIL-FAIXAS DE RENDA)
 Existência de filho tido nascido vivo no período de referência de 12 meses anteriores a 31/07/2010 — a variável dependente assume valor igual a um se a mulher teve filho nesse período, e zero caso contrário.

	0 filhos em 2009	1 filho em 2009	2 filhos em 2009	3 filhos em 2009	4 filhos em 2009
Renda per capita familiar	coef/robust SE	coef/robust SE	coef/robust SE	coef/robust SE	coef/robust SE
Até R\$ 70,00	0,0112* (0,0046)	0,0224*** (0,0016)	0,0106*** (0,0013)	0,0002 (0,0009)	0,0011 (0,0017)
Nº de observações	29.397	215.684	240.380	250.095	110.004
De R\$ 70,01 a R\$ 140,00	—	-0,0095*** (0,0027)	0,0141*** (0,0015)	-0,0014*** (0,0009)	-0,0015*** (0,0016)
Nº de observações	—	91.290	184.509	271.876	125.380
De R\$ 140,01 a R\$ 280,00	—	-0,0019** (0,0009)	0,0049*** (0,0007)	-0,0059*** (0,0008)	-0,0006 (0,0011)
Nº de observações	—	434.464	474.456	264.776	146.071
De R\$ 0,00 a R\$ 140,00	—	0,0131*** (0,0013)	0,0091*** (0,0009)	-0,0018*** (0,0006)	-0,0013*** (0,001)
Nº de observações	—	354.038	465.962	553.528	244.164†
De R\$ 0,00 a R\$ 280,00	—	0,0077*** (0,0008)	0,0005 (0,0005)	0,0022*** (0,0006)	-0,0006 (0,0008)
Nº de observações	—	821.205	1.100.665	654.778	390.226†

Notas: Erro padrão robusto entre parênteses. p -valor: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$. † indica a presença de desequilíbrio nas covariadas. As observações em **negrito** se referem às populações em que, após o corte dos valores extremos de Propensity Score, foi realizado o pareamento sem substituição, para, posteriormente, se estimar o efeito do tratamento. Nas demais, foram realizados apenas o corte dos valores extremos de p -score.

Em razão das disparidades regionais, da heterogeneidade cultural e das diferenças do número de beneficiários do Bolsa Família nas cinco regiões brasileiras (sendo que mais de 50% dos beneficiários residem na região nordeste), optou-se por realizar também a análise da fecundidade em nível regional, cujo resultado é apresentado na Tabela 3. Embora haja diferenças regionais, o comportamento em relação à fecundidade foi muito parecido em quase todas as regiões. Os maiores valores de ATT foram encontrados no grupo de mulheres que tinham apenas um filho em julho de 2009.

De acordo com os resultados obtidos, a região Centro-Oeste foi a que obteve o maior valor do efeito, apresentando um ATT de 4,65 pontos percentuais, ou seja, uma beneficiária do Bolsa Família tem probabilidade 4,65 pontos percentuais maior de gerar o segundo filho em comparação com o seu *match* do grupo de controle. Nessa mesma região, para a amostra de mulheres que tinham dois filhos em julho de 2009, o ATT foi de 1,03 pontos percentuais, sendo que após o terceiro filho o efeito do PBF sobre a fecundidade foi praticamente zero (e com baixa significância estatística).

As estimativas apontaram que a região Nordeste, que possui o maior número de beneficiárias, apresentou o segundo maior valor de ATT, cuja magnitude foi de 1,58 pontos percentuais para as mulheres com apenas um filho em 2009. Já para a amostra de mulheres que possuíam dois filhos, o sinal do impacto foi negativo, o que poderia indicar um desincentivo do PBF sobre a fecundidade caso a magnitude do ATT não fosse próxima de zero.

O valor de ATT para as mulheres com quatro filhos da região Sudeste deve ser desconsiderado, pois a amostra não apresentou equilíbrio nas covariadas. Além disso, essa população violou a hipótese de sobreposição. Para as demais amostras, todas estavam devidamente equilibradas nas covariadas e nenhuma violou a hipótese de sobreposição.

Por fim, em nível regional, o valor mais expressivo do efeito do PBF sobre a fecundidade foi encontrado na Região Centro-Oeste (4,65 pontos percentuais de diferença entre beneficiárias e não beneficiárias). No entanto, embora esse valor de ATT seja positivo e estatisticamente significativo, essa magnitude de efeito destoa dos demais resultados regionais. Sendo assim, esse valor de ATT apresenta baixa relevância econômica no que tange ao incentivo da fecundidade em relação às beneficiárias.

Em suma, o impacto em cada região foi semelhante ao nacional, ou seja, foi verificado ou a ausência de efeito do Programa sobre a fecundidade, ou valores de ATT muito baixos, pois mesmo que apresentem significância estatística, denotam baixa relevância econômica.

Em razão do trabalho de Berbel (2011), que aponta um incentivo do PBF sobre a fecundidade das beneficiárias com idade entre 16 e 19 anos, surge a necessidade de se verificar a fecundidade por faixas de idade. Assim, foi estimado o efeito médio do tratamento sobre as tratadas em quatro faixas etárias. A primeira faixa é justamente a faixa de idade utilizada por Berbel (2011), a segunda faixa é a mesma adotada pelo Sistema Único de Saúde (SUS) e as duas últimas são combinações das faixas etárias remanescentes das categorias de análise do SUS.²³

As estimativas por faixas de idade apontam algumas semelhanças com os resultados obtidos anteriormente, pois os maiores valores de ATT pertencem ao grupo de mulheres com apenas um filho em 2009. Para esse mesmo grupo, nas duas primeiras faixas etárias, observa-se que o efeito do tratamento foi de 5,25 e 2,98 pontos percentuais, respectivamente, indicando que o PBF propicia um pequeno incentivo à fecundidade, para as mulheres com apenas um filho. A partir da faixa etária dos 25 anos, esse incentivo praticamente desaparece.

Por fim, o valor positivo mais elevado de ATT foi observado na amostra de mulheres com apenas um filho e na primeira faixa de idade, ou seja, uma beneficiária do PBF com idade entre 16 e 19 anos tem probabilidade 5,25 pontos percentuais maior de gerar o segundo filho em comparação com o seu *match* do grupo de controle.

²³<http://tabnet.datasus.gov.br/cgi/idb1998/fqa05.htm>

**Tabela 3.** EFEITO DO TRATAMENTO SOBRE OS TRATADOS (REGIÕES DO BRASIL).

Existência de filho tido nascido vivo no período de referência de 12 meses anteriores a 31/07/2010 — a variável dependente assume valor igual a um se a mulher teve filho nesse período, e zero caso contrário.

Regiões	1 filho em 07/2009 coef/robust SE	2 filhos em 07/2009 coef/robust SE	3 filhos em 07/2009 coef/robust SE	4 filhos em 07/2009 coef/robust SE
NORTE	0,0107** (0,0047)	0,003 (0,003)	-0,0054*** (0,0019)	-0,0026 (0,0027)
Nº de observações	34.323	63.633	82.722	40.278
NORDESTE	0,0158*** (0,0012)	-0,0016** (0,0008)	0,0034*** (0,0007)	-0,0027** (0,0012)
Nº de observações	391.936	445.603	311.262	123.626
CENTRO-OESTE	0,0465*** (0,0089)	0,0103*** (0,0033)	0,0009 (0,0034)	-0,0063 (0,0191)
Nº de observações	10.739	33.016	37.705	12.328
SUDESTE	0,0072** (0,0035)	0,0047*** (0,0016)	-0,0036** (0,0017)	0,0053** (0,0026)
Nº de observações	58.669	141.306	106.983	65.560 [‡]
SUL	0,0140*** (0,0049)	0,0063* (0,0032)	0,0027 (0,0029)	0,0228 (0,0451)
Nº de observações	26.208	44.516	39.403	16.575

Notas: Erro padrão robusto entre parênteses. *p*-valor: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$. † indica a presença de desequilíbrio nas covariadas. As observações em **negrito** se referem às populações em que, após o corte dos valores extremos de *Propensity Score*, foi realizado o pareamento sem substituição, para, posteriormente, se estimar o efeito do tratamento. Nas demais, foram realizados apenas o corte dos valores extremos de *pscore*. As estimativas foram realizadas somente para as mulheres com renda per capita *ex-ante* de R\$ 0,00 até R\$ 140,00 (limite de elegibilidade ao PBF).

6.1. Análise de Sensibilidade

A análise de sensibilidade não exclui a possibilidade de que fatores não observáveis existam, mas questiona o “quanto de viés de seleção” deve estar presente para que seja possível aceitar a hipótese nula de ausência de efeito do tratamento. Por meio das informações da Tabela 5, se observa que a robustez ao viés de seleção varia entre os resultados das diferentes amostras utilizadas neste estudo. O efeito do tratamento parece ser robusto a uma possível presença de viés de seleção somente para as amostras das mulheres extremamente pobres, pois quanto maior o valor do nível crítico Γ , mais robusta a variável se apresenta.

Para as amostras das mulheres extremamente pobres (renda familiar per capita *ex-ante* de até R\$ 70,00), o nível crítico de Γ é de 1,20, indicando que se as variáveis não observáveis induzirem a uma diferença na razão de probabilidades de receber o tratamento entre as mulheres dos grupos de tratamento e controle por um fator de 1,20, pode-se questionar o efeito positivo das regras do PBF sobre a fecundidade das beneficiárias. De acordo com a Tabela 5, a fecundidade das mulheres que possuem apenas um filho em 2009 (com renda per capita de R\$ 70,01 a R\$ 140,00) também parece ser robusta à presença de viés de seleção.

No entanto, a fecundidade nas amostras regionais e por faixas etárias (com exceção das mulheres com idade de 16 a 19 anos) aparenta ser menos robusta à presença de variáveis não observáveis, dado que apresentam um valor crítico muito próximo da unidade. Os resultados apontam que em algumas amostras o nível crítico de Γ é igual a 1,10, ou seja, o viés de seleção capaz de gerar dúvidas acerca de um efeito positivo do PBF sobre a fecundidade ocorre quando variáveis não incluídas no modelo respondem

Tabela 4. EFEITO DO TRATAMENTO SOBRE OS TRATADOS (FAIXAS ETÁRIAS).

Existência de filho tido nascido vivo no período de referência de 12 meses anteriores a 31/07/2010 — a variável dependente assume valor igual a um se a mulher teve filho nesse período, e zero caso contrário.

Idade	1 filho em 07/2009 coef/robust SE	2 filhos em 07/2009 coef/robust SE	3 filhos em 07/2009 coef/robust SE	4 filhos em 07/2009 coef/robust SE
De 16 a 19 anos	0,0525*** (0,005)	-0,0171** (0,0079)	0,0145*** (0,0022)	-
Nº de observações	26.346	9.689	1.563	-
De 20 a 24 anos	0,0298*** (0,0029)	-0,0159*** (0,0019)	0,0009 (0,0031)	-0,0057 (0,0071)
Nº de observações	148.640	151.947	51.303	11.511
De 25 a 34 anos	0,0069*** (0,0016)	-0,0014*** (0,0008)	0,0032*** (0,0008)	-0,0007 (0,0014)
Nº de observações	277.465	516.472	347.894	142.535
De 35 a 50 anos	-0,0029 (0,0018)	0,0010* (0,0005)	0,0006 (0,0007)	-0,0021 (0,0012)
Nº de observações	98.131	356.303	230,177 [‡]	105,285 [‡]

Notas: Erro padrão robusto entre parênteses. *p*-valor: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$. [‡] indica a presença de desequilíbrio nas covariadas. Não havia observações suficientes para se calcular o ATT na amostra com faixa etária dos 16 aos 19 anos e quatro filhos. As observações em **negrito** se referem às populações em que, após o corte dos valores extremos de *Propensity Score*, foi realizado o pareamento sem substituição, para, posteriormente, se estimar o efeito do tratamento. Nas demais, foram realizados apenas o corte dos valores extremos de *pscore*. As estimativas foram realizadas somente para as mulheres com renda per capita *ex-ante* de R\$0,00 até R\$140,00 (limite de elegibilidade ao PBF).

por uma diferença de 10% na razão de probabilidades de participação no PBF entre as mulheres dos grupos de tratamento e controle.

Por fim, de acordo com DiPrete & Gangl (2004), os limites de Rosenbaum impõem os piores cenários possíveis dos resultados. Sendo assim, um Γ de 1,10 não implica necessariamente em ausência de efeito do PBF sobre a fecundidade, mas indica que o intervalo de confiança para o efeito do tratamento inclui zero se uma variável não observável causar uma diferença em 1,10 vezes na razão de probabilidades de participação ao tratamento.

7. CONSIDERAÇÕES FINAIS

Este trabalho se propôs a investigar um possível incentivo do Programa Bolsa Família ao aumento da fecundidade das beneficiárias, em decorrência do seu desenho pró-natalista. O diferencial deste estudo reside na construção de um contrafactual mais adequado ao problema de interesse, bem como na análise desse impacto em um maior período de exposição das beneficiárias aos efeitos do PBF, pois os estudos realizados até o momento avaliaram essa questão apenas no período inicial do Programa.

Nesse contexto, programas sociais podem ter impactos praticamente imediatos sobre a redução da pobreza, da miséria e da desigualdade, bem como sobre as taxas de matrículas escolares e a oferta de trabalho, mas não sobre a fecundidade. Faz-se necessário um maior período de exposição aos efeitos do Programa para que as beneficiárias assimilem os benefícios da transferência de renda e decidam se vale a pena, ou não, gerar mais um filho para ampliar a renda oriunda do Programa.

Dessa forma, controlando por variáveis observáveis e usando o método de *Propensity Score Matching*, com as técnicas de *matching* propostas por Imbens (2014), foi verificado que, em nível nacional, o efeito médio do tratamento sobre as tratadas (ATT) é de 1,31 pontos percentuais para as mulheres que tiveram

**Tabela 5.** Análise de sensibilidade para a fecundidade das beneficiárias do PBF

Amostras (Brasil)	Γ	p -crítico	Amostras (Brasil)	Γ	p -crítico
1 filho em 2009 e renda per capita de R\$ 0,00 a R\$ 70	1	< 0,0001	2 filhos em 2009 e renda per capita de R\$ 0,00 a R\$ 70	1	< 0,0001
	1,05	< 0,0001		1,05	0,000059
	1,10	0,001708		1,10	0,003329
	1,15	0,053961		1,15	0,051703
	1,20	0,365303		1,20	0,277467
1 filho em 2009 e renda per capita de R\$ 70,01 a R\$ 140	1	0,000012	2 filhos em 2009 e renda per capita de R\$ 70,01 a R\$ 140	1	0,199961
	1,05	0,000791		1,05	0,36873
	1,10	0,015895			
	1,15	0,118921			
1 filho em 2009 e renda per capita de R\$ 0,00 a R\$ 140	1	0,000632	2 filhos em 2009 e renda per capita de R\$ 0,00 a R\$ 140	1	0,000852
	1,05	0,105752		1,05	0,087346
				1,10	0,377528
Amostras (regiões)	Γ	p -crítico	Amostras (regiões)	Γ	p -crítico
1 filho em 2009 e Região Centro-Oeste	1	0,194816	2 filhos em 2009 e Região Centro-Oeste	1	0,456762
	1,05	0,321389		1,05	0,288455
1 filho em 2009 e Região Nordeste	1	0,002957	2 filhos em 2009 e Região Nordeste	1	0,000264
	1,05	0,193479		1,05	0,016888
				1,10	0,199703
1 filho em 2009 e Região Norte	1	0,360906	2 filhos em 2009 e Região Norte	1	0,472729
	1,05	0,148876		1,05	0,282201
1 filho em 2009 e Região Sudeste	1	0,184101	2 filhos em 2009 e Região Sudeste	1	0,078431
	1,05	0,501038		1,05	0,325982
1 filho em 2009 e Região Sul	1	0,109938	1 filho em 2009 e Região Sul	1	0,172143
	1,05	0,277283		1,05	0,371866
Amostras (faixas etárias)	Γ	p -crítico	Amostras (faixas etárias)	Γ	p -crítico
1 filho em 2009 e idade de 16 a 19 anos	1	0,023166	2 filhos em 2009 e idade de 16 a 19 anos	1	0,381028
	1,05	0,083677		1,05	0,265945
	1,10	0,212503			
1 filho em 2009 e idade de 20 a 24 anos	1	0,007291	2 filhos em 2009 e idade de 20 a 24 anos	1	0,016214
	1,05	0,157501		1,05	0,156137

Notas: A segunda coluna contém os valores de p^+ para os limites de Rosenbaum. Os valores de p -crítico em negrito se referem ao p^- para os limites de Rosenbaum. A ausência de viés de seleção devido a variáveis não observáveis ocorre quando $\Gamma = e^{\gamma} = 1$.

o segundo filho, ou seja, uma beneficiária do PBF tem probabilidade 1,31 pontos percentuais maior de gerar o segundo filho em comparação com o seu *match* do grupo de controle, a um nível de significância de 1%. Para as mulheres que tiveram o terceiro filho, o ATT foi de 0,91 pontos percentuais. No entanto, embora os valores de ATT sejam estatisticamente significantes, os seus valores em nível nacional são baixos, indicando baixa relevância econômica do impacto do PBF sobre a fecundidade das beneficiárias nesse âmbito.

Com o intuito de verificar a ocorrência de comportamentos diferentes daqueles encontrados em nível nacional, foram realizadas estimativas por regiões do Brasil (em razão da heterogeneidade cultural). Em geral, os resultados regionais apresentaram baixo valor de efeito médio do tratamento, sendo que as regiões Centro-Oeste e Nordeste apresentaram os maiores valores válidos de ATT. Nas estimativas por faixas etárias, os resultados obtidos são semelhantes aos resultados obtidos nos níveis nacional e regional, pois foi verificado um pequeno incentivo do PBF sobre a fecundidade para as mulheres que tiveram o segundo filho.

Dessa forma, embora não sejam expressivos, os resultados obtidos neste estudo demandam, por parte dos gestores de políticas públicas, um acompanhamento da fecundidade das beneficiárias, principalmente das que residem nas regiões Centro-Oeste e Nordeste (em especial para as faixas de idade de 16 a 19 anos e de 20 a 24 anos), haja vista que essas duas regiões representam aproximadamente 56,04% do total de beneficiários do Programa Bolsa Família.²⁴

Por fim, importante referir que este trabalho ficou limitado às informações oriundas dos filhos nascidos vivos, quando o ideal seria contabilizar também os filhos nascidos mortos e as gravidezes interrompidas ou mal sucedidas. Dessa forma, essa limitação pode acarretar numa subestimação do impacto do PBF sobre a fecundidade das beneficiárias. No entanto, mesmo que estejam subestimados, os resultados com sinal positivo foram consistentes e contrariaram os estudos anteriores, o que pode ter relação com o tempo de exposição aos efeitos do Programa.

Assim, para elucidar essa questão, ficam recomendadas novas pesquisas que possibilitem captar a intenção de engravidar das beneficiárias, contabilizando os filhos nascidos vivos e mortos e as gravidezes interrompidas. Alternativamente, ficam recomendados estudos que avaliem o impacto da Medida Provisória nº 535, de 2 de junho 2011, que alterou o limite de benefícios por filho, passando de 3 BV's para 5 BV's.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abadie, A., & G. Imbens. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics*, 29(1), 1–11.
- Abadie, A., & Imbens, G. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1), 235–267.
- Abadie, A., & Imbens, G. (2008). On the failure of the bootstrap for matching estimators. *Econometrica*, 76(6), 1537–1558.
- Abadie, A., & Imbens, G. (2012). *Matching on the estimated propensity score*. Disponível em: <http://www.hks.harvard.edu/fs/aabadie/pscore.pdf>
- Becker, G. S. (1960). An economic analysis of fertility. In A. Coale (Ed.), *Demographic and economic change in developed countries* (pp. 209–231). Princeton, NJ: Princeton University Press.
- Berbel, C. S. (2011). *Interações econômicas entre capital humano e fecundidade* (Tese de Doutorado em Economia, Fundação Getúlio Vargas–FGV, Rio de Janeiro). Disponível em: <http://hdl.handle.net/10438/9947>
- Crump, R. K., Hotz, V. J., Imbens, G. W. & Mitnik, O. A. (2008). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1), 187–199. doi: 10.1093/biomet/asn055
- DiPrete, T. A., & Gangl, M. (2004). Assessing bias in the estimation of causal effects : Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological*

²⁴https://www.beneficiosociais.caixa.gov.br/consulta/beneficio/04.01.00-00_00.asp



- Methodology*, 34(1), 271–310. doi: 10.1111/j.0081-1750.2004.00154.x
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161. Disponível em: <http://www.jstor.org/stable/1912352>
- Imbens, G. W. (2014, Mar). *Matching methods in practice: Three examples* (NBER Working Paper N° 19959). Cambridge, MA: National Bureau of Economic Research. doi: 10.3386/w19959
- Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5–86.
- Lalonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review*, 76(4), 604–620.
- Rocha, R. C. B. d. (2009, Dec 10). Programas condicionais de transferência de renda e fecundidade: Evidências do Bolsa Família [artigo apresentado em congresso]. In *31º Encontro Brasileiro de Econometria*, Foz do Iguaçu, Paraná. Disponível em: <http://bibliotecadigital.fgv.br/ocs/index.php/sbe/EBE09/paper/view/1104>
- Rosenbaum, P. R. (2002). Attributing to treatment in matched observational studies. *Journal of the American Statistical Association*, 97(457), 183–192.
- Rosenbaum, P. R. (2010). *Design of observational studies*. New York: Springer.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Signorini, B. A., & Queiroz, B. L. (2011). *The impact of Bolsa Família Program in the beneficiary fertility* (Texto para Discussão N° 439). Belo Horizonte, MG: Cedeplar/UFMG. Disponível em: <http://www.cedeplar.ufmg.br/pesquisas/td/TD%20439.pdf>
- Simões, P., & Soares, R. B. (2012). Efeitos do Programa Bolsa Família na fecundidade das beneficiárias. *Revista Brasileira de Economia*, 66(4), 445–468. doi: 10.1590/S0034-71402012000400004
- Souza, A. P., Oliveira, P. P., Duarte, J., Pereira, L. F. V. N. & Gadelha, S. R. B. (2013). *Uma análise dos determinantes da focalização do Programa Bolsa Família*. Disponível em: http://www.cepal.org/ofilac/noticias/paginas/9/49309/Brito_paper.pdf

APÊNDICE A. PROVA DAS PROPOSIÇÕES DO MODELO TEÓRICO

Primeiramente, é preciso definir que a derivada de BV em relação à n é igual a zero, quando n é igual a zero. Outras condições do modelo são:

$$\frac{\partial p}{\partial n} = p_n, \quad \frac{\partial BF}{\partial j} > 0, \quad \frac{\partial BF}{\partial k} > 0 \quad \text{e} \quad \frac{\partial U(0, Z)}{\partial n} = 0.$$

Prova da Proposição 1

As famílias desejam maximizar a sua utilidade sujeita à restrição orçamentária

$$\max U(n, Z) \quad \text{s.a} \quad p(n, \theta)n + \pi Z = R + BF(j, k, Y, \theta).$$

Montando o Lagrangeano,

$$L = U(n, Z) + \lambda [R + BF(j, k, Y, \theta) - p(n, \theta)n - \pi Z]$$

$$L_n = U_n = \lambda \left[p_n n + p(n, \theta) - BF_j(1 - \theta)\bar{j} - BF_k \theta \bar{k} + BF_Y \frac{R}{(n + c)^2} \right] \quad (\text{A-1})$$

$$L_Z = U_Z = \lambda \pi \quad (\text{A-2})$$

$$L_\lambda = p(n, \theta)n + \pi Z = R + BF(j, k, Y, \theta). \quad (\text{A-3})$$

Dividindo (A-1) por (A-2) e fazendo

$$\theta = BF_j(1 - \theta)\bar{j} + BF_K\theta\bar{k} - p(n, \theta) - BF_Y \frac{R}{(n + c)^2}$$

$$\frac{U_n}{U_Z} = \frac{np_n - \phi}{\pi}.$$

Para uma pequena variação de n , isto é, $\Delta n = 1$, por simplicidade, assume-se que o efeito em $BF_Y \approx 0$. Deste modo se chega à equação de n :

$$n = \frac{1}{p_n} \left(\frac{\pi U_n}{U_Z} + \phi \right). \quad (\text{A-4})$$

Para uma variação maior no número de filhos, não se pode desprezar o efeito em BF_Y , de tal modo que

$$\hat{\phi} = BF_j(1 - \theta)\bar{j} + BF_K\theta\bar{k} - p(n, \theta).$$

Então tem-se que

$$\frac{U_n \pi}{U_Z} + \hat{\phi} = np_n + BF_Y \frac{R}{(n + c)^2}.$$

Fazendo $\frac{U_n \pi}{U_Z} + \hat{\phi} = x$,

$$x = np_n + BF_Y \frac{R}{(n + c)^2}$$

$$x(n + c)^2 = n(n + c)^2 p_n + BF_Y R$$

$$(np_n - x)(n + c)^2 = -BF_Y R$$

$$(np_n - x)(n - c) = -BF_Y Y$$

$$n^2 p_n + nc - xn - xc = -BF_Y Y$$

$$n^2 p_n + n(c - x) + BF_Y Y = 0.$$

Solucionando a equação quadrática,

$$n = \frac{(x - c)}{2p_n} \pm \sqrt{(c - x)^2 - 4BF_Y Y}. \quad (\text{A-4}')$$

Note que em (A-4') n depende da renda per capita da família. Em outras palavras, não se consegue isolar o efeito dessa variável.

Inserindo essa equação em (A-3) tem-se

$$Z = \frac{1}{\pi} (R + BF - np). \quad \square \quad (\text{A-5})$$

Prova da Proposição 2

Inserindo a equação (A-1) na equação (A-2):

$$\frac{U_n}{U_Z} = \frac{np_n - \phi}{\pi}.$$

Note que se n aumenta p_n cai e ϕ aumenta. No limite, o custo de troca entre os dois bens é zero, mas quando n cresce fica mais caro em termos de utilidade trocar um filho pelo consumo de outros bens. \square



Prova da Proposição 3

Para esse caso, se deseja estimar o benefício marginal do nascimento do primeiro filho. Primeiramente, suponha que a família se enquadre na categoria de renda “extremamente pobre”. Nesse caso, deve-se supor que $n = n_{\min}$ e, tomando-se o diferencial da função de benefício do Programa Bolsa Família para um, $\Delta n = 1$, tem-se

$$BF^* \approx BF_{n_{\min}} + \Delta n \sum \frac{\partial BF}{\partial n}(n_{\min}).$$

Com exceção da variação da renda per capita, todas as derivadas da função de benefício são iguais a zero nesse ponto então teremos

$$BF^* \approx 77 + \Delta n \frac{\partial BF}{\partial Y} \frac{\partial Y}{\partial n}(n_{\min}).$$

E sabemos que

$$\frac{\partial BF}{\partial Y} \frac{\partial Y}{\partial n}(n_{\min}) = -\frac{\partial BF}{\partial Y} \frac{R}{(n+c)^2} = -\frac{\partial BF}{\partial Y} \frac{R}{c^2}.$$

Então, teremos:

$$BF^* \approx 77 - \frac{\partial BF}{\partial Y} \frac{R}{c^2}.$$

Por construção (decorrente do grupo de renda familiar) sabemos que

$$77 > \frac{\partial BF}{\partial Y} \frac{R}{c^2}.$$

Em outras palavras, o primeiro filho reduz a renda per capita da família. Sendo a família da categoria dos “extremamente pobres”, sua renda cairá ainda mais, corroborando o resultado acima. Formalmente temos que

$$BF^* > 0.$$

□

Prova da Proposição 4

Suponha que o número de filhos da família está quase no limite de n_{\max} . Nesse caso, pode-se dizer que

$$\lim_{n \rightarrow n_{\max}} \phi = BF_j(1 - \theta)n_{\max} + BF_K \theta n_{\max} - p(n_{\max}, \theta) - BF_Y \frac{R}{(n_{\max} + c)^2}.$$

A equação acima pode ser escrita do seguinte modo:

$$\lim_{n \rightarrow n_{\max}} \phi = -p(n_{\max}, \theta) + \sum \frac{\partial BF}{\partial n}(n_{\max}).$$

Note que o termo $\sum \frac{\partial BF}{\partial n}(n_{\max})$ tende a zero e a função de custos do nascimento de filhos decresce à medida que $n \rightarrow n_{\max}$, ou seja, marginalmente o acréscimo de benefício tende a zero dado o nascimento de mais filhos. Desta forma, a escolha da família pelo nascimento de uma nova criança não é mais relacionada com a renda advinda do programa. □

APÊNDICE B. TABELAS

Tabela B-1. ESTATÍSTICAS DESCRITIVAS DA VARIÁVEL DE INTERESSE.

Existência de filho tido nascido vivo no período de referência de 12 meses anteriores a 31/07/2010 — a variável dependente assume valor igual a um se a mulher teve filho nesse período, e zero caso contrário.

Nº filhos em 07/2009	PBF = 1			PBF = 0		
	observações	média	desvio padrão	observações	média	desvio padrão
Mulheres com renda per capita <i>ex-ante</i> de R\$ 0,00 a R\$ 70,00						
Zero	13.236	0,1757	0,3806	16.161	0,2036	0,4027
Um	97.960	0,1516	0,3587	117.724	0,1287	0,3349
Dois	108.823	0,0994	0,2992	131.557	0,0824	0,2749
Três	174.969	0,0622	0,2415	118.747	0,0599	0,2374
Quatro	82.385	0,0746	0,2628	46.864	0,0754	0,264
Mulheres com renda per capita <i>ex-ante</i> de R\$ 70,01 a R\$ 140,00						
Um	42.423	0,1874	0,3902	48.867	0,2083	0,4061
Dois	84.325	0,1205	0,3255	100.184	0,1196	0,3245
Três	140.396	0,0609	0,2392	131.480	0,0541	0,2262
Quatro	67.291	0,0587	0,2351	58.089	0,0587	0,2352
Mulheres com renda per capita <i>ex-ante</i> de R\$ 0,00 a R\$ 140,00						
Um	161.907	0,1644	0,3706	192.131	0,1549	0,3618
Dois	211.755	0,1015	0,3019	254.207	0,0949	0,2931
Três	345.838	0,0579	0,2336	286.431	0,0536	0,2253
Quatro	151.582	0,0651	0,2468	108.199	0,0657	0,2478
Região	PBF = 1			PBF = 0		
	observações	média	desvio padrão	observações	média	desvio padrão
Mulheres com um filho em julho de 2009 e renda per capita <i>ex-ante</i> de R\$ 0,00 a R\$ 140,00						
Norte	16.008	0,2443	0,4297	18.315	0,2391	0,4265
Nordeste	200.780	0,1691	0,3748	191.156	0,1799	0,3842
Centro-Oeste	4.810	0,2692	0,4436	5.929	0,2282	0,4197
Sudeste	25.871	0,2132	0,4096	32.798	0,2261	0,4183
Sul	11.621	0,1832	0,3868	14.587	0,1848	0,3882
Mulheres com dois filhos em julho de 2009 e renda per capita <i>ex-ante</i> de R\$ 0,00 a R\$ 140,00						
Norte	29.201	0,1358	0,3426	34.432	0,1299	0,3362
Nordeste	341.631	0,0660	0,2483	184.918	0,0732	0,2605
Centro-Oeste	15.077	0,0768	0,2663	17.939	0,0881	0,2834
Sudeste	70.369	0,0936	0,2913	150.550	0,0906	0,2870
Sul	20.311	0,1040	0,3053	24.205	0,1129	0,3166

Fonte: Elaboração própria a partir dos dados do Censo 2010.

Notas: O valor da mediana foi 0,00 para todas as amostras. Em razão de a variável “teve filho” ser binária, os valores mínimos e máximos de cada amostra são, respectivamente, 0 e 1. As estatísticas descritivas foram feitas utilizando-se os pesos do Censo 2010.



Tabela B-2. Verificação da hipótese de sobreposição entre os grupos de controle e tratamento (faixas de renda).

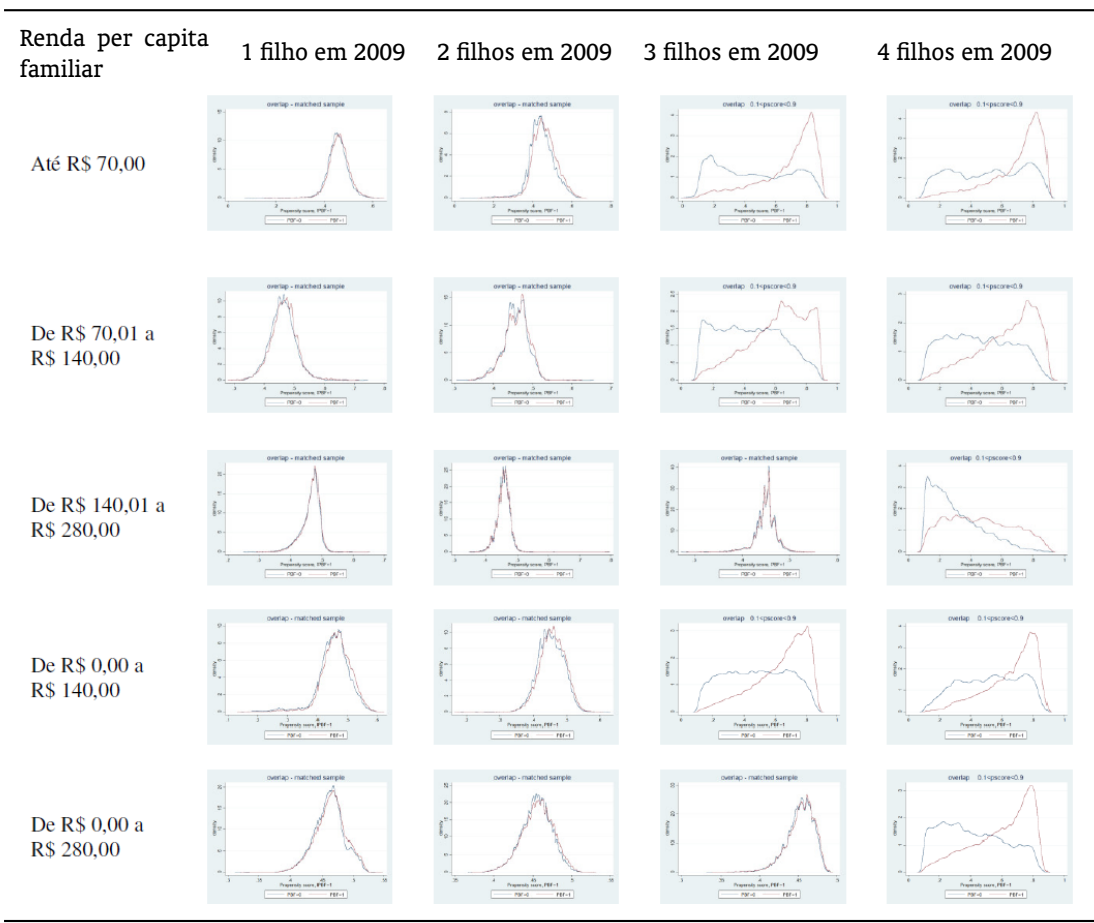


Tabela B-3. Pseudorresultado: análise da fecundidade em uma variável pré-tratamento (Brasil).

Renda per capita familiar	0 filhos em 2009 coef/robust SE	1 filho em 2009 coef/robust SE	2 filhos em 2009 coef/robust SE	3 filhos em 2009 coef/robust SE	4 filhos em 2009 coef/robust SE
Até R\$70,00	-	0,0353 *** (0,0019)	0,0356 *** (0,0027)	0,0574 *** (0,0034)	0,0019 (0,0179)
Nº de observações	-	215.684	240.380	250.095	110.004
De R\$70,01 até R\$140,00	-	0,0286 *** (0,0029)	0,0056 (0,0036)	0,0174 *** (0,0038)	-0,0059 (0,0162)
Nº de observações	-	91.290	184.509	271.876	125.380
De R\$0,00 até R\$140,00	-	0,0364 *** (0,0013)	0,0200 *** (0,0022)	0,0522 *** (0,0025)	0,0209 *** (0,0042)
Nº de observações	-	354.038	465.962	553.528	244.164
De R\$140,01 até R\$280,00	-	0,0145 *** (0,0014)	0,0141 *** (0,0019)	-0,0035 (0,0032)	-0,0442 *** (0,006)
Nº de observações	-	434.464	474.456	264.776	146.071
De R\$0,00 até R\$280,00	-	0,0202 *** (0,0009)	0,0012 *** (0,0002)	-0,0154 *** (0,0022)	0,0198 *** (0,0035)
Nº de observações	-	821.205	1.100.665	654.778	390.226

Notas: Erro padrão robusto entre parênteses. *p*-valor: *** *p* < 0,01, ** *p* < 0,05, * *p* < 0,1. Coeficientes em **negrito** indicam que a hipótese de independência condicional é plausível para aquela amostra.



Regiões e Especialização na Agricultura Cafeeira: São Paulo no Início do Século XX*

Renato Perim Colistete†

Sumário: 1. Introdução; 2. Regiões e Expansão Cafeeira; 3. A Estrutura da Produção Agrícola; 4. Especialização Regional; 5. Especialização das Propriedades Rurais; 6. Pequenas e Grandes Propriedades na Oferta Agrícola; 7. Conclusões; Apêndice. Classificação das regiões de São Paulo e seus municípios.

Palavras-chave: Agricultura Cafeeira, Especialização Agrícola, São Paulo.

Códigos JEL: N53, O13, Q12.

Este artigo analisa a especialização da produção agrícola nas regiões e nas propriedades rurais do estado de São Paulo no início do século XX, após décadas de transformações causadas pela difusão da lavoura do café. O trabalho utiliza dados de mais de 40 mil propriedades rurais para examinar a estrutura de produção e a especialização produtiva em onze regiões que cobrem a totalidade do território do estado de São Paulo no auge da economia cafeeira. Além das regiões, o artigo aborda como as propriedades rurais distribuíram seus recursos produtivos entre produto de exportação (café) e produtos destinados ao consumo interno. A análise traz à luz novas evidências, corroborando parte da literatura, mas divergindo de outras ideias estabelecidas na historiografia. No início do século XX, virtualmente todas as regiões de São Paulo produziam café, mas quase todas elas se especializaram em mais de um produto agrícola, com a exceção da região da Mogiana. Além das grandes propriedades ressaltadas na literatura, as pequenas propriedades também foram atraídas para a produção de exportação e se especializaram na produção do café. Por sua vez, as grandes propriedades não só se especializaram na produção de alimentos (além do café), mas dominaram a oferta dos principais produtos destinados ao mercado doméstico.

This article deals with agricultural specialization in regions and farms of the state of São Paulo in the early twentieth century, after several decades of changes caused by the coffee boom. We use data from more than forty thousand farms to examine the agrarian structure and specialization in eleven regions which cover the whole state of São Paulo at the time. The article also shows how farms chose between export (coffee) and internal market crops. The results lend support to part of the literature

*Agradeço aos participantes do 10º Congresso Brasileiro de História Econômica, em Juiz de Fora; do 42º Encontro Nacional de Economia, em Natal; e especialmente a Maria Lúcia Lamounier e Pedro Garcia Duarte pelos comentários e sugestões a uma versão preliminar deste trabalho.

†Departamento de Economia, Universidade de São Paulo. Rua Luciano Gualberto, 908, Cidade Universitária, São Paulo-SP. CEP 05508-010. E-mail: rcolistete@usp.br



but add new evidence which diverges from established notions among historians. In the early twentieth century, all regions of São Paulo cultivated coffee, but nearly all of them specialized in more than one agricultural product, except the Mogiana region. Besides the plantations highlighted by the literature, small farms were also drawn into export production and specialized in coffee. In turn, large properties not only specialized (apart from coffee) in food crops, but dominated the supply of the main products for internal markets.

1. INTRODUÇÃO

No início do século XX, São Paulo havia alcançado a posição de maior produtor de café no Brasil e no mundo, resultado da expansão da fronteira agrícola que ganhou novo impulso em meados do século anterior. Longe de homogênea e linear, a expansão da fronteira agrícola assumiu diferentes características de acordo com o momento em que as regiões foram incorporadas ao *boom* cafeeiro e as condições fisiográficas, demográficas e institucionais existentes. O deslocamento de sitiantes e fazendeiros em busca de novas terras, a entrada de imigrantes europeus e a chegada das ferrovias no interior de São Paulo refletiram as possibilidades criadas para a produção de exportação tanto em terras já há tempos exploradas com atividades agrícolas diversas e criação de animais quanto em áreas cobertas por florestas, campos e cerrados, habitadas por indígenas e por poucas famílias pioneiras, geralmente dispersas e isoladas (Pestana, 1923, cap.1–3; Monbeig, 1998).

Dois aspectos da diversidade da expansão agrícola em São Paulo foram ressaltados por geógrafos e historiadores já nas primeiras décadas do século XX. Primeiro, o fato do avanço do café ter sido desigual entre as diferentes áreas da província (e, depois, estado), com regiões e municípios empregando parcelas crescentes de suas terras cultivadas na agricultura de exportação enquanto outros o faziam apenas marginalmente ou de forma secundária. Na realidade, as regiões marginais da cafeicultura também iriam fazer parte do *boom* exportador, mas por meio da especialização em atividades que se beneficiaram com o aumento da renda interna e da demanda derivada do café. Segundo, a expansão cafeeira em diversas regiões do estado de São Paulo combinou-se com o crescimento simultâneo da produção agrícola destinada ao mercado doméstico. Produtos de exportação e produtos para o consumo interno foram cultivados lado a lado mesmo nas regiões em que o café mais se expandiu.¹

Este artigo analisa a especialização da produção agrícola no estado de São Paulo no início do século XX, enfatizando a variedade da estrutura agrária encontrada em suas regiões e propriedades rurais. Os dados utilizados referem-se a mais de 40 mil propriedades rurais pesquisadas pela Secretaria de Agricultura de São Paulo no ano agrícola de 1904–1905 (São Paulo, 1906-1907). Inicialmente, o artigo examina de que maneira diferentes regiões adaptaram suas estruturas agrárias à rápida expansão dos mercados no século XIX e qual foi o perfil de especialização produtiva daí resultante. A análise utiliza evidências e medidas quantitativas que demonstram que a especialização agrícola variou substancialmente entre as regiões do estado de São Paulo, inclusive entre as próprias zonas tipicamente cafeeiras, resultado que corrobora as pesquisas de historiadores e geógrafos mencionados antes. No início do século XX havia três zonas especializadas na produção cafeeira, mas apenas uma delas (a Mogiana) o era exclusivamente no café. As outras duas zonas cafeeiras (Araraquarense e Paulista) e as demais regiões de São Paulo especializavam-se em outros produtos além do café, geralmente voltados ao consumo doméstico.²

A outra questão tratada no artigo diz respeito à esfera das decisões das propriedades rurais: qual foi a combinação de culturas agrícolas adotada por produtores rurais sob diferentes circunstâncias geográficas e econômicas? Em que medida o domínio crescente do café levou também a uma especialização

¹Entre os principais trabalhos do início do século XX com essa perspectiva, ver Deffontaines (1936a, 1936b); Denis (1909); James (1932); Milliet (1939/1982b); Monbeig (1937); Platt (1935); Prado Jr. (1945b); Sampaio (1900–1901).

²Ver o Apêndice para a classificação das regiões de São Paulo e seus municípios.

no nível microeconômico das propriedades rurais? Pequenas e grandes propriedades distinguiram-se em suas estratégias de produção? Como veremos, as evidências indicam que, no início do século XX, as grandes propriedades do estado de São Paulo tenderam a ser as mais especializadas no cultivo do café, mas as pequenas propriedades acompanharam essa tendência apresentando elevados níveis de especialização no produto de exportação nas zonas tipicamente cafeeiras. Esse resultado põe em xeque a noção de que a lavoura do café em São Paulo tenha sido uma atividade essencialmente vinculada a grandes fazendas e elevadas economias de escala.³ Por extensão, o resultado conflita também com a tese de que o Brasil divergiria de outros países da América Latina, tais como Venezuela, Costa Rica e partes da Colômbia, cuja produção cafeeira contou com uma substancial, ou mesmo majoritária, presença de pequenas propriedades.⁴

Por outro lado, as grandes propriedades de São Paulo apresentaram maior flexibilidade nas suas decisões de produção do que geralmente é assumido: embora as pequenas unidades rurais em geral fossem mais especializadas nos produtos para consumo doméstico (tais como milho e feijão), as grandes propriedades apresentaram também alto grau de especialização nesse tipo de produção. Mais importante ainda, em termos de volume de produção, as grandes propriedades dominaram a oferta dos principais produtos agrícolas destinados ao consumo doméstico, contrariando a percepção comum de que as pequenas propriedades tenham tradicionalmente se especializado e se tornado as principais fornecedoras para o mercado interno.⁵ A produção de alimentos somente foi dominada por pequenas propriedades em certas regiões (a Capital e o Vale do Paraíba e Litoral Norte), mas não em uma escala suficiente para alterar o predomínio das grandes propriedades na oferta de mantimentos no conjunto do estado de São Paulo. Essas questões serão tratadas em detalhe nas seções seguintes.

2. REGIÕES E EXPANSÃO CAFEIRA

Desde a primeira metade do século XIX, o crescimento do café incorporou rapidamente novas áreas à produção de exportação e, por volta de 1900, as regiões que cultivavam o produto eram substancialmente distintas da área — Vale do Paraíba e Litoral Norte — onde se iniciou a expansão cafeeira em São Paulo. Ao lado de regiões ocupadas nos primórdios da colonização, antigas produtoras de cana-de-açúcar, alimentos e gado (como a Capital, o Vale do Paraíba e Sorocaba) e outras na fronteira da expansão do café (Ribeirão Preto), havia áreas extensas que ainda no início do século XX eram pouco conhecidas e habitadas majoritariamente por populações indígenas e famílias pioneiras vivendo em áreas isoladas (como nas regiões dos rios do Peixe, Aguapé e Paranapanema).⁶ As condições fisiográficas, o regime de propriedade da terra e o tipo de atividade econômica existentes influenciaram diretamente o modo com que a demanda internacional do café foi assimilada pelas estruturas de produção locais.

Não por acaso, os primeiros estudos sobre a expansão cafeeira em São Paulo deram atenção especial aos fatores que deveriam ser utilizados para classificar as regiões que constituíam a extensa e diversificada área agrícola do estado. Ainda na década de 1930, Pierre Deffontaines constatou a dificuldade de se estabelecer uma classificação coerente das regiões de São Paulo tendo em vista a instabilidade da exploração agrícola e da população, que rapidamente deslocava-se em busca de novas terras e oportunidades, enquanto novos produtos e atividades produtivas sucediam-se na mesma região. Considerando arbitrária a classificação já então convencional das áreas ocupadas pelas linhas ferroviárias, Deffontaines propôs uma tipologia baseada nas características geológicas e topográficas do estado de São Paulo, chegando a um total de 19 regiões fisiográficas (Deffontaines, 1936a, 1936b).

³Uma formulação clássica da visão dual que opõe grandes e pequenas propriedades na economia de exportação é a de Prado Jr. (1945/1969, pp.163–4 e 248–52). Ver também Camargo (1952, p.191) e Costa (1966/1998, pp.75 e 522). Críticos de Caio Prado Jr. também compartilham da mesma visão: ver uma análise explícita (para o Rio de Janeiro) nesses termos em Castro (1987, pp.22–61) e de Castro (1988).

⁴Ver, por exemplo, Yarrington (1997, p.64) e Solbrig (2008, p.508).

⁵Para essa interpretação no período aqui tratado, ver: Cano (1981, pp.60–61) e Font (1990, pp.18–21).

⁶Ver Sampaio (1890); Cobra (1923); Prado Jr. (1945b, pp.924–31); Monbeig (1998, pp.183–88); e Monbeig (1957).



Apesar das qualificações de Deffontaines, outras tipologias continuaram enfatizando aspectos econômicos e populacionais de São Paulo no século XIX e início do século XX, deixando em segundo plano as características da geografia física. O estudo mais conhecido e influente nessa linha foi o de Sérgio Milliet (1939/1982b). Utilizando critérios relacionados à expansão da cultura do café, aos movimentos populacionais e às áreas de influência das ferrovias, Milliet dividiu o estado de São Paulo em sete regiões: zona Norte (inclusive litoral), Central e aquelas abrangendo municípios tributários das ferrovias Mogiana, Paulista, Araraquarense, Noroeste e Sorocabana (Milliet, 1939/1982b, pp.13–15). Em outro estudo, o autor incluiu três zonas adicionais à sua classificação anterior: Capital–Santos, Baixa Sorocabana e Litoral Sul (Milliet, 1939/1982a, pp.70–71).

Posteriormente, outros autores adotaram classificações com critérios alternativos. Pierre Monbeig (1945-1946), por exemplo, reelaborou a proposta original de Deffontaines, combinando elementos geográficos e econômicos em uma nova classificação com quatro regiões fisiográficas básicas, doze sub-regiões secundárias e trinta e cinco zonas definidas de acordo com características econômicas aproximadamente homogêneas dos municípios. Talvez pela complexidade das subdivisões ou pelo predomínio ainda de critérios fisiográficos, a classificação de Monbeig acabou recebendo pouca acolhida. Já a classificação de Milliet passou a ser amplamente utilizada nas décadas seguintes. O trabalho mais conhecido que empregou a divisão regional de Milliet foi o de José Francisco de Camargo sobre a população e a economia de São Paulo, já com a inclusão das três zonas às sete originais de *Roteiros do Café* e algumas adaptações marginais (Camargo, 1952, pp.23–48). Mais tarde, Joseph Love adotou a classificação de Milliet adaptada por Camargo (Love, 1982, 41–51). Thomas Holloway, por sua vez, também utilizou a tipologia de Milliet, mas com modificações em relação à adaptação de Camargo (Holloway, 1974, Appendix I).

Na análise a seguir, utilizaremos a divisão regional de Milliet na versão modificada de Holloway, por parecer mais adequada ao período aqui estudado. As principais diferenças da versão de Holloway em relação à de Camargo são a manutenção da zona da Paulista de acordo com a divisão original de Milliet (apenas integrando a Alta Paulista à zona Noroeste, em vez da divisão “Paulista”, “Alta Paulista” e “Baixa Paulista”) e a criação de uma nova região em torno de Sorocaba, desmembrada da zona da Capital (Holloway, 1974, Appendix I). As onze zonas e seus respectivos municípios são apresentados no Apêndice, com suas datas de fundação e nomes oficiais em 1905. Para facilitar a identificação, as denominações atuais dos municípios são incluídas entre parênteses. Ao todo, são 171 municípios em 1905 que produziam aproximadamente 70% do café exportado pelo Brasil.⁷

Na época, as áreas de produção de café mais antigas já estavam em pleno declínio, enquanto outras regiões ainda se encontravam no início da expansão cafeeira. Milliet demonstrou que a situação econômica dos municípios de São Paulo não estava associada apenas ao sucesso da cultura do café, que pela sua natureza extensiva tendia a ser efêmera, mas dependia em especial da capacidade das localidades diversificarem sua produção agrícola e suas atividades urbanas de produção de manufaturas, de comércio e financiamento. Assim, enquanto o Vale do Paraíba era caracterizado como “zona morta, que o café desbravou, povoou, enriqueceu e abandonou antes que criasse raízes para o progresso”, a Mogiana e a Paulista tiveram destino diferente, pois nelas “a decadência do café teve por consequência não o abandono das terras, como na zona norte, mas o ensaio de outras culturas substitutivas”, além da produção de manufaturas e serviços. Como observa Milliet, citando um município particular, “de como tais fatores podem modificar as condições do abandono cafeeiro, vê-se pelo aumento contínuo da população de Rio Claro em oposição à decadência da produção de café” (Milliet, 1939/1982b, pp.37 e 50–52).

O cenário de expansão e declínio identificado por Milliet pode ser resumido por meio da evolução da população das onze zonas tributárias descritas no Apêndice, considerando os anos de 1872, 1886, 1907 e 1920 (Tabela 1).⁸

⁷ Calculado com dados de exportação para 1910, extraídos de Brasil (1917, p.117).

⁸ Milliet (1939/1982b) utilizou dados de 1836, 1854, 1886, 1920 e 1935.

Tabela 1. População por Regiões do Estado de São Paulo, 1872–1920.

Regiões	1872	%	1886	%	1907	%	1920	%
Capital	46.775	5,6	74.895	6,2	352.375	11,3	654.578	14,3
Vale do Paraíba e Litoral Norte	251.423	29,9	335.932	27,7	474.919	15,3	484.699	10,6
Sorocaba	78.763	9,4	103.789	8,6	150.532	4,8	190.423	4,2
Central	141.234	16,8	179.021	13,0	394.476	12,7	487.175	10,6
Mogiana	122.057	14,5	184.021	15,2	621.782	20,0	811.974	17,7
Paulista	69.614	8,3	133.733	11,0	362.109	11,7	530.257	11,6
Araraquarense	16.347	1,9	33.151	2,7	281.905	9,1	583.801	12,7
Noroeste	–	–	–	–	21.248	0,7	151.771	3,3
Alta Sorocabana	40.377	4,8	81.670	6,7	219.979	7,1	392.006	8,6
Baixa Sorocabana	42.131	5,0	64.847	5,4	114.748	3,7	134.227	2,9
Santos e Litoral Sul	32.300	3,8	42.434	3,5	113.468	3,7	161.950	3,5
Total Estado de São Paulo	841.021	100,0	1.233.493	100,0	3.107.541	100,0	4.582.861	100,0

Nota: Até 1887, o município de Bauru (Noroeste) pertencia a Lençóis.

Fontes: Brasil (1876), São Paulo (1888), Brasil (1917), Brasil (1926).

A Tabela 1 mostra que a zona do Vale do Paraíba e Litoral Norte sofreu acentuado declínio em sua participação na população total do estado de São Paulo entre 1872 e 1920, resultado da decadência do café e da pequena expansão relativa das atividades que o substituíram na maior parte dos municípios. Outras regiões antigas — Central e Sorocaba — também tiveram reduções relativas da população, embora vários de seus municípios tenham sido capazes de diversificar suas atividades econômicas. A Mogiana e a Paulista, prolongamentos da região Central na expansão cafeeira do final do século XIX, já diminuía o ritmo do crescimento populacional nas primeiras décadas do século XX diante das mais recentes fronteiras agrícolas — Araraquarense, Noroeste e Alta Sorocabana. A maior parte dessas três últimas zonas era ainda formada, em 1907, por extensas áreas que seriam colonizadas nos anos seguintes. A estrutura da produção agrícola de todas as regiões é analisada a seguir.

3. A ESTRUTURA DA PRODUÇÃO AGRÍCOLA

As informações quantitativas sobre a atividade agrícola no início do século XX demonstram de forma clara as diferentes situações das regiões de São Paulo após décadas de expansão cafeeira. Os dados utilizados nesta e nas seções seguintes foram extraídos da *Estatística Agrícola e Zootécnica do Estado de São Paulo*, elaborada pela Secretaria da Agricultura, Comércio e Obras Públicas do Estado de São Paulo e que coletou informações detalhadas das propriedades rurais de 171 municípios e mais de 50 mil propriedades rurais em 1904–1905 (São Paulo, 1906-1907).⁹ A Tabela 2 apresenta o número de municípios e propriedades com informações disponíveis para cada uma das regiões descritas no Apêndice e no estado de São Paulo como um todo.¹⁰

Conforme mostra a Tabela 3, a zona da Mogiana possuía a maior extensão de terras particulares, com cerca de 17,8% da área total ocupada por propriedades rurais no estado de São Paulo (coluna “Área Total”). Não por coincidência, as outras zonas tributárias das ferrovias também se destacavam em termos de área das propriedades rurais: Alta Sorocabana (14,1%), Araraquarense (12,6%) e Paulista (12,1%). As menores áreas de propriedades rurais no conjunto do estado encontravam-se, como esperado, nas regiões de Santos e Litoral Sul (1,8%) e Capital (1,8%).

⁹Os dados foram extraídos da edição crítica da *Estatística Agrícola* realizada por Bassanezi & Francisco (2003).

¹⁰Seis municípios não tiveram suas informações publicadas (conf. Bassanezi & Francisco, 2003): Apiaí, Iguape, Itú, Mogi das Cruzes, Santa Isabel e Tietê. Por esse motivo, a Tabela 2 totaliza os dados disponíveis de 165 municípios e suas propriedades.

**Tabela 2.** Municípios e Propriedades Rurais, São Paulo, 1904-1905.

Regiões	Nº de municípios	Nº de propriedades
Capital	7	2.524
Vale do Paraíba e Litoral Norte	35	10.921
Sorocaba	13	2.570
Central	18	7.344
Mogiana	27	8.086
Paulista	17	4.563
Araraquarense	16	5.597
Noroeste	1	341
Alta Sorocabana	16	4.015
Baixa Sorocabana	11	2.553
Santos e Litoral Sul	4	758
Total Estado de São Paulo	165	49.272

Fonte: São Paulo (1906-1907).

Tabela 3. Participação das Regiões na Área Agrícola das Propriedades Rurais, São Paulo, 1904-1905 (em %).

Regiões	Área Cultivada	Área em Mata	Área em Capoeira	Área em Campos e Pastos	Área em Brejos e Terra Imprestável	Área Total
Capital	1,1	1,2	2,2	2,3	1,8	1,8
Vale do Paraíba e Litoral Norte	12,4	10,6	17,0	5,7	27,1	11,0
Sorocaba	2,9	1,7	12,4	7,4	5,0	6,3
Central	15,8	4,5	12,7	8,2	7,1	8,8
Mogiana	24,9	12,0	9,6	28,0	4,4	17,8
Paulista	16,1	9,7	6,4	17,2	11,4	12,1
Araraquarense	13,4	19,6	7,0	8,4	8,0	12,6
Noroeste	0,4	12,9	0,4	0,4	0,6	4,7
Alta Sorocabana	7,6	17,2	12,1	14,7	14,8	14,1
Baixa Sorocabana	4,8	6,9	18,2	7,7	17,9	9,5
Santos e Litoral Sul	0,6	3,6	2,0	0,0	2,2	1,8
Total Estado de São Paulo	100,0	100,0	100,0	100,0	100,0	100,0

Notas: Dados originais em alqueire paulista, daqui em diante citado apenas como alqueire. 1 alqueire = 2,42 hectares ou 24.200 metros quadrados. Dados agregados de 49.272 propriedades agrícolas, conforme Tabela 2.

Fonte: São Paulo (1906-1907).

Um quadro distinto é obtido quando se considera a área efetivamente empregada na produção agrícola. Novamente a Mogiana aparece com a maior área cultivada (24,9%) do total do estado de São Paulo, mas após a Paulista (16,1%) a antiga região Central (15,8%) aparece com destaque, seguida pela Araraquarense (13,4%) e outra zona antiga na expansão agrícola paulista, o Vale do Paraíba e Litoral Norte (12,4%). Alta Sorocabana (7,6%), Baixa Sorocabana (4,8%) e Noroeste (0,4%) apresentavam uma proporção de área cultivada substancialmente menor do que suas participações no total da área ocupada, o que indica que em tais regiões a produção agrícola mais intensiva ainda se encontrava em seus primórdios (Tabela 3).

De fato, a intensidade da exploração da terra variou significativamente entre as regiões do estado de São Paulo, conforme é possível constatar na Tabela 4. No conjunto do estado, apenas 12,4% da área das propriedades rurais eram cultivadas. A Central era a região com a maior área cultivada (22,2%), da mesma forma que a sua área com mata ocupava uma das menores parcelas (17,5%) das propriedades, indicando a expansão pioneira do café na região e a destruição acelerada das matas nativas (Dean, 1997, cap.5). As outras zonas que possuíam participações relativamente altas (considerando a média do estado) da área cultivada incluíam Mogiana (17,4%), Paulista (16,4%), Vale do Paraíba e Litoral Norte (14,0%) e Araraquarense (13,2%). Destas, apenas a região da Araraquarense conservava uma proporção ainda significativa de matas nativas (53,5%), situação que seria transformada nas décadas seguintes com o rápido crescimento demográfico e a ocupação agrícola (Monbeig, 1998, pp.181–95).

A agricultura de São Paulo no início do século XX foi dominada pelo principal produto de exportação: nada menos do que 55,3% da área cultivada eram utilizados na lavoura do café. Do ponto de vista das regiões, porém, as diferenças eram enormes. Enquanto a Mogiana (69,0%), Araraquarense (63,8%), Paulista (61,9%) e Alta Sorocabana (55,2%) registravam bem mais da metade de suas terras cultivadas produzindo café, outras regiões apresentavam participações ínfimas, tais como a zona da Capital (1,1%), a Baixa Sorocabana (4,8%) e a região de Sorocaba (8,7%) (Tabela 5).

Outros produtos de destaque eram destinados ao mercado interno, em especial milho e feijão, com 25,5% e 9,8% da área cultivada, respectivamente, no total do estado de São Paulo. Esses eram produtos de amplo consumo, podendo ser cultivados intercalados entre as fileiras dos cafeeiros ou em áreas separadas.¹¹ Também nesses casos a variação regional era expressiva. Baixa Sorocabana (77,4%), Sorocaba

Tabela 4. Distribuição da Área Agrícola das Propriedades Rurais por Regiões, São Paulo, 1904-1905 (em %).

Regiões	Área Cultivada	Área em Mata	Área em Capoeira	Área em Campos e Pastos	Área em Brejos e Terra Imprestável	Área Total
Capital	7,9	24,3	26,5	39,5	1,9	100,0
Vale do Paraíba e Litoral Norte	14,0	33,0	33,1	15,8	4,1	100,0
Sorocaba	5,4	14,0	42,7	36,3	1,6	100,0
Central	22,2	17,5	30,6	28,2	1,4	100,0
Mogiana	17,4	23,1	11,6	47,6	0,5	100,0
Paulista	16,4	27,5	11,4	43,0	1,7	100,0
Araraquarense	13,2	53,5	12,0	20,2	1,2	100,0
Noroeste	1,2	94,1	1,7	2,9	0,2	100,0
Alta Sorocabana	6,8	41,1	18,2	32,0	1,9	100,0
Baixa Sorocabana	6,2	25,0	40,9	24,6	3,4	100,0
Santos e Litoral Sul	4,3	69,5	23,8	0,1	2,2	100,0
Total Estado de São Paulo	12,4	34,3	21,3	30,3	1,8	100,0

Fonte: São Paulo (1906-1907).

¹¹Von Delden Laerne (1885), pp.241 e 246–7. Sobre o amplo consumo de milho no planalto paulista, Holanda (1956/1994).

**Tabela 5.** Participação da Área Plantada com Produtos na Área Cultivada, São Paulo, 1904-1905 (em %).

Regiões	Café	Cana-de-Açúcar	Algodão	Arroz	Milho	Feijão	Diversas
Capital	1,1	4,0	0,7	2,0	42,5	23,1	26,7
Vale do Paraíba e Litoral Norte	43,8	6,2	0,2	3,9	24,3	16,6	5,0
Sorocaba	8,7	2,5	12,0	4,8	54,3	15,0	2,7
Central	49,8	4,7	0,5	2,3	28,3	11,1	3,3
Mogiana	69	1,9	0,0	3,4	16,9	8,6	0,2
Paulista	61,9	1,9	0,0	2,9	22,9	10,1	0,3
Araraquarense	63,8	2,5	0,1	6,0	22,4	4,6	0,5
Noroeste	44,7	1,5	0,0	5,5	41,1	7,0	0,2
Alta Sorocabana	55,2	2,0	0,5	4,1	31,6	5,5	1,2
Baixa Sorocabana	4,8	1,6	1,5	8,3	68,1	9,3	6,3
Santos e Litoral Sul	12,9	1,3	0,0	35,9	7,6	4,3	38,0
Total Estado de São Paulo	55,3	2,9	0,5	3,9	25,5	9,8	2,1

Fonte: São Paulo (1906-1907).

(69,3%) e Capital (65,6%) ocupavam mais da metade de sua área cultivada com milho e feijão. As regiões em que predominava o café também possuíam uma área substancial plantada com milho e feijão. Outros produtos para consumo doméstico reunidos na categoria Diversas — tais como mandioca, batata, frutas, hortaliças, fumo e alfafa — tinham um peso significativo particularmente na área cultivada de regiões com pouca expressão na produção de café, como Santos e Litoral Sul (38,0%) e Capital (26,7%). A cana-de-açúcar era comum em praticamente todas as regiões, enquanto algodão e arroz ocupavam a maior área das propriedades de algumas zonas específicas — Sorocaba (algodão) e Baixa Sorocabana e Santos e Litoral Sul (arroz) (Tabela 5).

Também na Tabela 5 é possível notar que as regiões que empregavam maiores parcelas da terra no produto de exportação eram distintas daquelas que ocupavam maiores extensões de sua área cultivada com produtos típicos do mercado interno. Esse é um ponto que será analisado detalhadamente na próxima seção, mas aqui já é possível identificar algumas tendências gerais. A observação dessa tendência pode ser auxiliada por uma análise de correlação entre a área cultivada com café e a área com produtos destinados ao mercado doméstico nas diferentes regiões de São Paulo. Os coeficientes da Tabela 6 mostram que a maior presença do café era associada com menor participação da área cultivada com produtos do mercado interno. Todos os coeficientes relacionando café e tais produtos são negativos (conforme a primeira coluna da Tabela 6), com exceção da cana-de-açúcar, cujo coeficiente é positivo mas bastante baixo (0,0109). Vale lembrar que tais resultados referem-se à participação relativa dos produtos na área

Tabela 6. Correlação da Área Cultivada nas Regiões, São Paulo, 1904-1905.

Produtos	Café	Milho	Feijão	Arroz	Cana-de-Açúcar	Algodão	Diversas
Café	1	–	–	–	–	–	–
Milho	–0,5743	1	–	–	–	–	–
Feijão	–0,4631	0,3479	1	–	–	–	–
Arroz	–0,3549	–0,3724	–0,4317	1	–	–	–
Cana-de-açúcar	0,0109	–0,0745	0,6528	–0,3838	1	–	–
Algodão	–0,4431	0,5053	0,2932	–0,1036	–0,0434	1	–
Diversas	–0,6423	–0,2176	0,1787	0,7466	–0,033	–0,1237	1

Nota: Correlações de Pearson com os dados da Tabela 5.

Fonte: São Paulo (1906-1907).

cultivada dentro de cada região, ou seja, os coeficientes refletem o resultado agregado das decisões dos proprietários rurais quanto à distribuição da área plantada entre diferentes lavouras. Naturalmente, regiões com áreas cultivadas mais extensas — a Mogiana, por exemplo — tiveram influência bem maior nos resultados da área cultivada total dos produtos (como visto na Tabela 3).

Realmente, considerando agora os produtos específicos, as zonas da Mogiana (33,4%), Paulista (21,7%), Araraquarense (13,9%) e Central (13,3%) possuíam as maiores participações na área cultivada com café, totalizando 82,3% do estado de São Paulo em 1904-1905 (ver Tabela 7). Além disso, o que chama a atenção é que essas mesmas zonas estavam também entre as que possuíam as maiores áreas dos principais produtos destinados ao abastecimento interno — milho, feijão e arroz. Ou seja, mesmo com menos alqueires de sua área cultivada destinados a alimentos do que ao café, como foi visto antes, o tamanho absoluto da área cultivada com produtos para o mercado doméstico garantia um lugar proeminente às zonas cafeeiras também na lavoura de mantimentos em São Paulo. Por exemplo, Mogiana (17,8%), Paulista (17,4%) e Central (16,4) somavam as maiores áreas de produção de milho, seguidas pela antiga zona cafeeira do Vale do Paraíba e Litoral Norte (11,1%) e a Araraquarense (10,6%). Mogiana (23,5%) e Paulista (19,9%) também concentravam as maiores áreas com lavouras de feijão, embora aqui a zona do Vale do Paraíba e Litoral Norte (19,8%) tivesse uma posição de destaque, acompanhada pela região Central (16,7%).

Da mesma forma, a Mogiana (23,3%) concentrava a maior área de cultivo de arroz, acompanhada pela Araraquarense (18,6%), Paulista (14,4%) e o Vale do Paraíba e Litoral Norte (11,7%). Um perfil similar pode ser identificado em relação à área cultivada com cana-de-açúcar. Apenas no caso de algodão e de produtos diversos é que outras regiões ganhavam destaque: Sorocaba possuía 55,8% da área cultivada com algodão, enquanto Vale do Paraíba e Litoral Norte (28,5%), Central (23,9%) e a Capital (13,4%) somavam as maiores áreas de culturas diversas, boa parte formada por alimentos, o que indica a importância dessas zonas para o abastecimento dos mercados locais.

Esses resultados revelam características importantes da estrutura produtiva agrícola de São Paulo no início do século XX. Se, de um lado, o café destacava-se como produto principal, por outro lado, outros produtos agrícolas ocupavam uma parcela significativa da terra cultivada. As variações regionais foram grandes, com algumas zonas concentrando-se em produtos não destinados à exportação. Mas as regiões cafeeiras também eram aquelas com as maiores áreas absolutas dedicadas a produtos tipicamente do mercado interno, a exemplo do milho, feijão e arroz. Ainda assim, outros produtos importantes para o abastecimento local tinham a sua maior área de cultivo localizada nas zonas que não se destacavam na produção de café na época. Essas características aparentemente contraditórias serão analisadas com

Tabela 7. Participação das Regiões na Área Cultivada, São Paulo, 1904-1905 (em %).

Regiões	Café	Cana-de-Açúcar	Algodão	Arroz	Milho	Feijão	Diversas
Capital	0,02	1,4	1,6	0,5	1,7	2,4	13,4
Vale do Paraíba e Litoral Norte	9,2	24,5	4,0	11,7	11,1	19,8	28,5
Sorocaba	0,5	2,4	55,8	3,1	5,9	4,0	3,7
Central	13,3	24,0	15,5	9,0	16,4	16,7	23,9
Mogiana	33,4	17,7	0,3	23,3	17,8	23,5	3,0
Paulista	21,7	12,3	0,7	14,4	17,4	19,9	2,9
Araraquarense	13,9	10,4	2,7	18,6	10,6	5,7	3,0
Noroeste	0,4	0,2	0,0	0,7	0,8	0,4	0,1
Alta Sorocabana	7,3	5,0	7,7	7,9	8,9	4,1	4,3
Baixa Sorocabana	0,3	1,9	11,7	7,3	9,2	3,3	10,5
Santos e Litoral Sul	0,1	0,2	0,0	3,4	0,1	0,2	6,8
Total Estado de São Paulo	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Fonte: São Paulo (1906-1907).



mais detalhe na próxima seção.

4. ESPECIALIZAÇÃO REGIONAL

A especialização produtiva na agricultura de São Paulo no início do século XX pode ser medida por um índice que relacione a área destinada a uma lavoura específica em um região e a área ocupada por essa lavoura no estado como um todo. Um índice maior do que 1 indica que a região é especializada no cultivo de um determinado produto.¹² Formalmente, o índice de especialização é definido como

$$esp = \frac{e_i/e}{E_i/E}, \quad (1)$$

onde esp é o índice de especialização agrícola, e_i é a área cultivada com o produto i na região, e é a área cultivada total na região, E_i é a área cultivada com o produto i no estado de São Paulo, e E é a área cultivada total no estado de São Paulo.

A Tabela 8 traz estimativas do índice de especialização agrícola regional de São Paulo no início do século XX. É possível ver inicialmente que as zonas da Mogiana, Paulista e Araraquarense distinguem-se pela especialização na produção de café. A Mogiana, que possuía a maior área cultivada com café do estado, era na época a única região que se especializava exclusivamente na produção cafeeira, enquanto a Paulista também se especializava na produção de milho e feijão e a Araraquarense no cultivo de arroz, além de café.

As demais regiões especializaram-se em outros produtos, ainda que todas elas também produzissem café. Mesmo sendo a quarta região com a maior área plantada com cafeeiros, conforme visto anteriormente, a zona Central não era na época especializada na cafeicultura, devido à diversificação de sua área cultivada com produtos como cana-de-açúcar, milho, feijão e diversos. A mais antiga região do café, o Vale do Paraíba e Litoral Norte, também já destinava relativamente mais terras para as lavouras de cana-de-açúcar, feijão e diversas, da mesma forma que a Capital prosseguiu em sua tradicional especialização na agricultura não-exportadora e, sobretudo, em produtos como batata, mandioca, hortaliças e

Tabela 8. Especialização Agrícola por Região segundo a Área Cultivada, São Paulo, 1904-1905.

Regiões	Café	Cana-de-Açúcar	Algodão	Arroz	Milho	Feijão	Diversas
Capital	0,01843	1,24404	1,39216	0,46583	1,53946	2,1721	11,975
Vale do Paraíba e Litoral Norte	0,73761	1,95995	0,31684	0,93649	0,89207	1,58553	2,27988
Sorocaba	0,17162	0,87546	20,42268	1,11997	2,15287	1,4608	1,33732
Central	0,83155	1,49808	0,97197	0,56152	1,02757	1,04645	1,49506
Mogiana	1,32513	0,70382	0,01203	0,92659	0,70772	0,93362	0,11998
Paulista	1,33459	0,75894	0,04506	0,88942	1,07534	1,22854	0,1779
Araraquarense	1,02537	0,76836	0,19563	1,37526	0,78465	0,41941	0,22148
Noroeste	0,87909	0,54188	0	1,55265	1,75734	0,78018	0,11484
Alta Sorocabana	0,97011	0,66758	1,02063	1,0492	1,17511	0,55017	0,55983
Baixa Sorocabana	0,06195	0,39433	2,43673	1,52031	1,90469	0,67588	2,18537
Santos e Litoral Sul	0,13749	0,25738	0	5,4934	0,1763	0,25956	10,93256

Nota: Especialização da produção agrícola medida pelo índice definido no texto. Um índice maior do que 1,0 indica especialização na produção do bem.

Fonte: São Paulo (1906-1907).

¹²O índice de especialização corresponde ao coeficiente de localização usado em várias disciplinas. Ver Nakamura & Paul (2009, pp.306-7).

forragem para animais, classificados como diversas. Outra região antiga, Santos e Litoral Sul, apresentou altos índices de especialização no cultivo de arroz e produtos diversos, igualmente reproduzindo um perfil já estabelecido no século anterior. Sorocaba, por sua vez, era a região mais especializada no cultivo do algodão, ao lado de produtos tradicionais (milho, arroz, feijão e diversos). As demais zonas, Noroeste, Alta Sorocabana e Baixa Sorocabana, as duas primeiras ainda em um estágio inicial de entrada de imigrantes e expansão da fronteira agrícola, apresentavam um perfil próximo ao das zonas antigas no que diz respeito à especialização agrícola, com destaque para a produção de algodão, arroz, milho e diversos (Tabela 8).

O fato de apenas a região da Mogiana ter sido exclusivamente especializada em café no início do século XX sugere que a diversificação agrícola era muito maior do que tende a ser reconhecida pelas análises agregadas da economia de São Paulo durante a expansão cafeeira. A maior parte das regiões cultivava uma variedade de produtos ao lado do café. Ainda assim, em que medida a convivência do café com culturas diversas era de fato expressiva? E mais ainda, seria correto assumir que a presença do café estava associada a uma menor diversificação da produção agrícola? Analogamente à especialização, o grau de diversificação da área cultivada nas regiões de São Paulo pode ser estimado com um índice específico, conforme a seguinte equação (Singh & Dhillon, 1984, pp.220–21):

$$\text{índice de diversificação agrícola} = 1 - \frac{\text{Percentual da área total cultivada com produtos } n}{\text{Número de produtos } n}, \quad (2)$$

onde “produtos n ” correspondem àqueles com 5% ou mais de área cultivada na região.

A Tabela 9 reproduz os índices de diversificação agrícola das regiões de São Paulo no início do século XX. As regiões com os maiores índices e, portanto, com a maior diversificação da produção agrícola eram a Capital, Santos e Litoral Sul e Sorocaba. Nenhuma delas era especializada no cultivo do café, embora todas apresentassem áreas plantadas com cafeeiros. As zonas com menor diversificação da produção agrícola compreendiam as três com especialização na lavoura cafeeira — Paulista, Mogiana e Araraquarense. Assim, o índice sugere que, de fato, a maior importância do cultivo do café era associada com menor diversificação agrícola nas regiões.

Esse resultado fica mais claro ao compararmos a área plantada com café e o índice de diversificação agrícola das regiões, como mostra a Figura 1. Há uma nítida relação entre aumento da participação da lavoura de café no total da área cultivada e diminuição do índice de diversificação agrícola nas regiões, inclusive naquelas que não eram especializadas na produção cafeeira, como por exemplo a zona Noroeste.

Tabela 9. Diversificação Agrícola, São Paulo, 1904-1905.

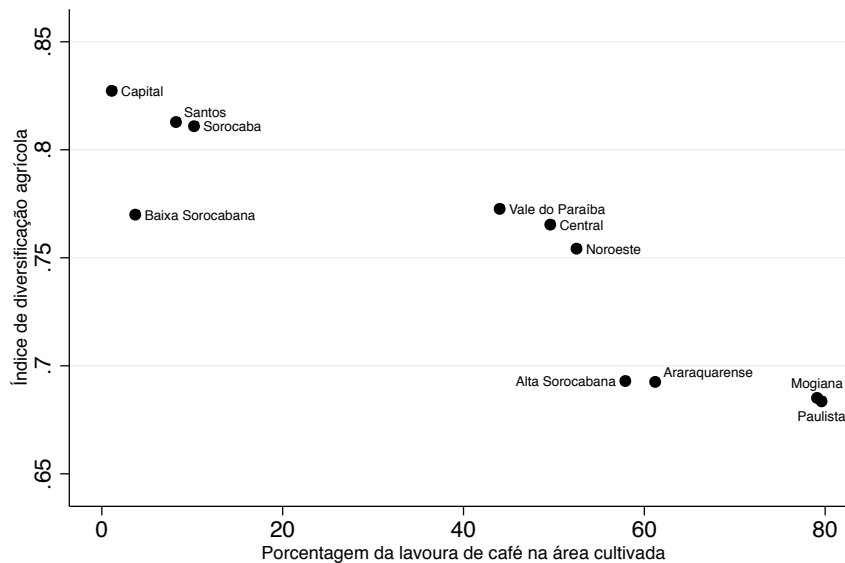
Regiões	Índice de diversificação agrícola
Capital	0,827
Vale do Paraíba e Litoral Norte	0,773
Sorocaba	0,811
Central	0,765
Mogiana	0,685
Paulista	0,684
Araraquarense	0,693
Noroeste	0,754
Alta Sorocabana	0,693
Baixa Sorocabana	0,770
Santos e Litoral Sul	0,813

Nota: Índice de diversificação agrícola conforme definido na equação (2).

Fonte: São Paulo (1906-1907).



Figura 1. Área Cultivada com Café e Diversificação Agrícola, São Paulo, 1904-1905.



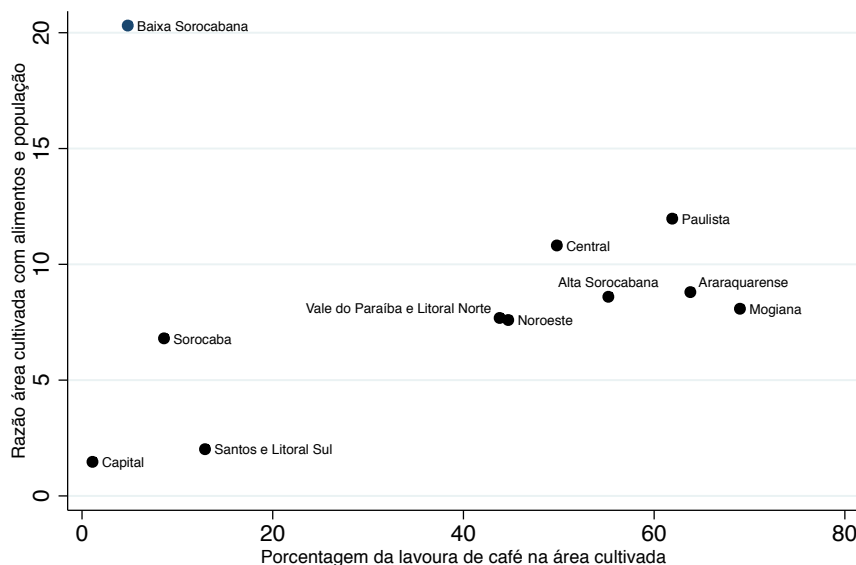
Fonte: Ver Tabela 2. Índice de diversificação definido na equação (2).

As regiões com os maiores índices de diversificação (Capital, Santos e Sorocaba) especializavam-se em produtos como milho, arroz e feijão.

É preciso observar, porém, que a queda da diversificação agrícola com o aumento da área plantada com café não implicava que a produção de outros produtos fosse inexpressiva nas regiões cafeeiras. Ao contrário, como foi visto antes, as zonas especializadas em café eram também as que possuíam as maiores áreas cultivadas no estado de São Paulo com os principais produtos para abastecimento interno, isto é, arroz, feijão e milho (ver Tabela 7). Há indícios inclusive de que as propriedades rurais nas zonas cafeeiras típicas tenham produzido mantimentos em quantidade suficiente para atender seus moradores e comercializar parte do produto nas localidades urbanas próximas. Na ausência de dados sobre o comércio entre municípios, uma estimativa da importância relativa da produção de alimentos pode ser feita comparando-a com a população residente em cada região.

A Figura 2 mostra que a área cultivada com alimentos em relação à população residente nas regiões aumenta conforme cresce a proporção da área ocupada com café no total da área cultivada. Essa relação pode ser interpretada como um indicador da capacidade de sustentação da população local com a produção de mantimentos, de forma que quanto maior for a razão café/área cultivada, maior a tendência de que os alimentos produzidos possam ter sido utilizados para autoconsumo e, eventualmente, para venda fora da propriedade rural. Tanto os proprietários das terras quanto os colonos imigrantes e nacionais podem ter sido responsáveis por essas lavouras. A única exceção à tendência verificada nas regiões foi a Baixa Sorocabana, que apresentou uma reduzida proporção de terras ocupadas com o café mas uma elevada área de produção de alimentos em relação à sua população, o que sugere que essa região pode ter produzido um excedente substancial para o mercado.

A análise agregada da produção agrícola revela, assim, um quadro variado de especialização produtiva entre as diferentes regiões do estado de São Paulo no início do século XX. A especialização exclusiva no cultivo do café limitou-se à maior região exportadora na época. Outras zonas cafeeiras novas, assim como as antigas, possuíam estruturas agrárias mais diversificadas, apesar da importância central do café nas suas economias. Não por coincidência, portanto, a maior área cultivada com produtos-chave para os mercados locais localizava-se com frequência justamente nas regiões que eram também especializadas

Figura 2. Importância da Produção de Alimentos, São Paulo, 1904-1905.

Fonte: Ver Tabela 2. Dados de população referentes ao ano de 1907.

na produção de café. Resta saber, porém, como se dava a alocação das terras cultivadas no âmbito das propriedades rurais individuais.

5. ESPECIALIZAÇÃO DAS PROPRIEDADES RURAIS

Um outro aspecto da especialização agrícola na economia cafeeira diz respeito às decisões das propriedades rurais — o que e quanto cultivar, dadas a situação dos mercados e as condições de solo, oferta de trabalho, transporte, entre outros fatores. Quais foram as estratégias de produção dos proprietários rurais nas regiões cafeeiras e nas zonas especializadas em outros produtos? Em particular, de que forma pequenas e grandes unidades produtivas distribuíam suas terras entre produtos de exportação e mercado interno?

Um passo inicial para responder a essas perguntas consiste em estabelecer critérios para a divisão das unidades produtivas em pequenas e grandes propriedades rurais. Adotaremos aqui a classificação utilizada por Prado Jr. (1945a, p.693), em que as propriedades até 25 alqueires foram definidas como pequenas, as maiores que 25 a 100 alqueires como médias e as maiores que 100 alqueires como grandes propriedades. Alguns anos depois, Sérgio Milliet acrescentou mais uma categoria à classificação de Caio Prado Jr. a fim de contemplar diferenças entre as grandes propriedades. Milliet definiu as grandes propriedades como aquelas maiores que 100 até 500 alqueires e uma nova classe — latifúndio — para as unidades maiores que 500 alqueires, critério que também é adotado a seguir (Milliet, 1939/1982a, p.70).¹³ Apesar da arbitrariedade implícita no critério, a classificação torna possível organizar as informações das propriedades rurais sob um princípio uniforme, permitindo compará-las entre as regiões.¹⁴

A Tabela 10 apresenta a participação da área plantada com café no total da área cultivada das propriedades rurais em São Paulo no início do século XX, segundo as classes de tamanho das propriedades

¹³A classificação de Caio Prado Jr. é apresentada e discutida com mais detalhe em Colistete & Lamounier (2014, pp.12–13).

¹⁴Como observou José Francisco Camargo, “No caso do Estado de São Paulo, por exemplo, uma propriedade de 100 alqueires pode, conforme a zona em que se situe, representar uma propriedade pequena, média ou grande”; mesmo assim, argumenta o autor adiante, tal tipo de classificação é “ordinariamente a base de medida mais usual e compreensiva.” (Camargo, 1952, pp.189–90).



agrícolas definidas por Caio Prado Jr. e Sérgio Milliet.

Os dados referem-se a informações individuais de 48.013 propriedades rurais pesquisadas pela *Estatística Agrícola* de 1904-1905. Os ajustes realizados nos dados originais consistiram principalmente em excluir as propriedades cujas informações sobre área e produção eram incompletas ou inexistentes.

Os dados da tabela indicam que, em São Paulo, propriedades maiores tendiam a possuir áreas relativamente mais extensas de terras cultivadas com café, principalmente quando as pequenas propriedades são comparadas com as outras classes de tamanho. No conjunto do estado, as pequenas propriedades destinavam 36,1% de sua área cultivada para a produção de café, contra 52,6% das médias propriedades, 51,3% das grandes propriedades e 55,3% dos latifúndios. A exceção notável a essa tendência ocorreu na região da Araraquarense, em que as pequenas propriedades dedicaram mais terras ao café (52,1%) do que as grandes propriedades rurais e latifúndios (20,6% e 46,5%, respectivamente).¹⁵ A alocação de terras para o café também era elevada (em relação à média estadual) entre as pequenas propriedades de outras duas regiões tipicamente cafeeiras (Mogiana, 58,3%; Paulista, 43,1%), além da Alta Sorocabana (44%). Esses números demonstram que o produto de exportação típico da época era atraente também para os pequenos proprietários e que a concentração no cultivo do café nas maiores regiões produtoras abrangia todas as classes de produtores rurais.

Já as médias propriedades (52,6%) tinham uma participação marginalmente mais alta do café na área cultivada do que as grandes propriedades (51,3%) no total do estado de São Paulo (Tabela 10). Esse resultado geral foi determinado pela elevada participação do café nas médias propriedades da Araraquarense (64,2%), muito superior à das grandes propriedades (20,6%) e dos latifúndios (46,5%). Considerando o estado como um todo, os latifúndios possuíam a maior área relativa plantada com cafeeiros (55,3%), embora a diferença com relação às médias (52,6%) e grandes (51,3%) propriedades tenha sido marginal. Além disso, em quatro regiões — inclusive a maior produtora, a Mogiana, além da Central, Capital e Alta Sorocabana — os latifúndios possuíam relativamente menos terras dedicadas ao café do que as grandes

Tabela 10. Participação do Café na Área Cultivada por Tamanho da Propriedade, São Paulo, 1904-1905.

Regiões	Pequena Propriedade	Média Propriedade	Grande Propriedade	Latifúndio
Capital	1,7	0,6	1,1	0,0
Vale do Paraíba e Litoral Norte	22,6	45,1	58,6	58,7
Sorocaba	7,3	7,7	9,0	11,8
Central	27,5	43,9	64,3	41,4
Mogiana	58,3	70,0	73,9	64,7
Paulista	43,1	55,0	66,0	66,5
Araraquarense	52,1	64,2	20,6	46,5
Noroeste	24,3	30,2	48,1	63,8
Alta Sorocabana	44,0	56,8	68,9	59,8
Baixa Sorocabana	4,8	6,7	7,1	2,6
Santos e Litoral Sul	2,0	6,7	12,6	58,3
Estado de São Paulo	36,1	52,6	51,3	55,3

Notas: Cálculo baseado nos dados individuais de 48.013 propriedades rurais do estado de São Paulo. Tamanhos das propriedades de acordo com a classificação de Prado Jr. (1945a) e Milliet (1939/1982a): i) pequenas propriedades: de 0 a 25 alqueires; ii) médias propriedades: de 25 a 100 alqueires; iii) grandes propriedades: de 101 e 500 alqueires, e iv) latifúndios: maiores que 500 alqueires. 1 alqueire = 2,42 hectares ou 24.200 metros quadrados.

Fonte: São Paulo (1906-1907).

¹⁵As pequenas propriedades da região da Capital também tiveram maior participação de café na área cultivada, mas essa proporção era ínfima (1,7%, conf. Tabela 10).

propriedades; e em outras três (Central, Araraquarense e Capital), menos área cultivada com café do que as médias ou pequenas propriedades. Essas figuras são uma clara evidência de que o cultivo do café não se expandiu linearmente com o tamanho da propriedade e que houve variações regionais significativas em comparação com os resultados agregados do conjunto do estado de São Paulo.

É possível também medir a atração que o café exercia para um amplo espectro de produtores agrícolas estimando o grau de especialização de cada propriedade individual segundo sua classificação por tamanho. Para tanto, utilizamos o mesmo índice de especialização calculado na seção anterior (conf. Tabela 8), mas agora aplicado às propriedades individuais. A Tabela 11 atesta que, na média do estado de São Paulo, os maiores índices de especialização na produção cafeeira eram encontrados nas grandes propriedades, seguidas pelos latifúndios. Em três zonas específicas, os latifúndios apareciam como as unidades mais especializadas no café — Mogiana com 71,1% e Noroeste com 43,8%, além de um produtor marginal, Santos e Litoral Sul, com 9,1%. No entanto, as maiores diferenças na especialização das grandes propriedades em relação às médias e pequenas ocorriam nas antigas regiões do café, o Vale do Paraíba e a Central. Nelas, grandes propriedades e latifúndios tinham um percentual de especialização igual ou superior à média do estado, enquanto as pequenas e médias apresentavam percentuais significativamente menores que os registrados para o estado como um todo (ver Tabela 11). Na ausência de dados comparativos para décadas anteriores, o domínio das propriedades maiores possui duas explicações possíveis: seja porque, na sua época áurea, as grandes propriedades e os latifúndios das antigas regiões eram mais especializados do que os seus similares nas novas zonas do café ou, o que parece mais provável, o declínio da cafeicultura nas regiões antigas levou a uma maior diversificação produtiva das pequenas e médias propriedades.

No que diz respeito às outras classes de tamanho, mais uma vez a Araraquarense se distinguiu pelo maior peso relativo (45,9%) das médias propriedades especializadas em café do que as grandes propriedades (43,6%) e os latifúndios (43,8%), da mesma forma que ocorria com as médias propriedades (42,4%) da Alta Sorocabana em relação aos latifúndios (38,9%). Por sua vez, 24,9% das pequenas propriedades eram especializadas na produção de café no estado. Acompanhando a tendência das outras classes de

Tabela 11. Propriedades Especializadas em Café segundo Classes de Tamanho, São Paulo, 1904-1905.

Regiões	Pequena Propriedade	Média Propriedade	Grande Propriedade	Latifúndio
Capital	0,1	0,3	1,1	0,0
Vale do Paraíba e Litoral Norte	14,7	31,4	47,1	43,9
Sorocaba	4,4	2,7	4,8	4,8
Central	23,5	25,5	56,4	39,6
Mogiana	51,3	66,1	68,7	71,1
Paulista	31,6	46,6	60,3	49,5
Araraquarense	38,0	45,9	43,6	43,8
Noroeste	13,9	20,0	31,0	43,8
Alta Sorocabana	31,8	42,4	50,7	38,9
Baixa Sorocabana	2,7	2,9	4,4	1,9
Santos e Litoral Sul	1,1	1,2	3,6	9,1
Estado de São Paulo	24,9	35,8	47,4	40,3

Notas: Especialização da produção agrícola medida pelo índice definido no texto (ver equação na página 396). Cálculo baseado nos dados individuais de 22.051 propriedades rurais com área cultivada com café no estado de São Paulo. Tamanhos das propriedades de acordo com a classificação de Prado Jr. (1945a) e Milliet (1939/1982a): i) pequenas propriedades: de 0 a 25 alqueires; ii) médias propriedades: de 25 a 100 alqueires; iii) grandes propriedades: de 101 e 500 alqueires, e iv) latifúndios: maiores que 500 alqueires. 1 alqueire = 2,42 hectares ou 24.200 metros quadrados. *Fonte:* São Paulo (1906-1907).



propriedades, esse percentual subia acentuadamente entre as pequenas unidades agrícolas das regiões tipicamente cafeeiras: 31,6% no caso da Paulista, 38,0% na Araraquarense e 51,3% na Mogiana, a maior zona produtora no início do século XX (Tabela 11).

Assim, embora as grandes propriedades tendessem a ser mais especializadas na produção de café, as pequenas propriedades também apresentavam elevados níveis de especialização na produção cafeeira, sobretudo nas zonas de expansão do café. Tal fato não é surpreendente dada a capacidade da lavoura do café adaptar-se a diferentes tamanhos de propriedade e regimes de trabalho (Gudmundson, 1989, p.221; Roseberry, 1995, pp.5–7; Topik & Clarence-Smith, 2003, pp.386–94). Uma implicação é que a ideia amplamente difundida de que a produção de café em São Paulo foi realizada basicamente em grandes fazendas, com alta capitalização e elevadas economias de escala — antítese do que ocorria, por exemplo, na Venezuela, Costa Rica e em partes da Colômbia —, deve ser qualificada.¹⁶ O mesmo se aplica às análises que associam pequenas unidades produtivas no Brasil fundamentalmente ao mercado interno, constituindo um mundo oposto ao domínio da exportação e das grandes propriedades.¹⁷ As pequenas e médias propriedades acompanharam a tendência de especialização regional e também foram atraídas para a produção de exportação, tendo em algumas zonas produtoras superado as grandes propriedades ou os latifúndios em seu grau de especialização na produção do café.

Uma questão complementar é saber qual era o perfil daquelas unidades que produziam bens tradicionalmente destinados ao consumo interno. A Tabela 12 apresenta a participação relativa das propriedades (por classes de tamanho) especializadas na produção de um produto típico de consumo interno, o milho, que era utilizado como alimento da população humana e dos animais de criação.

Tabela 12. Propriedades Especializadas em Milho segundo Classes de Tamanho, São Paulo, 1904-1905.

Regiões	Pequena Propriedade	Média Propriedade	Grande Propriedade	Latifúndio
Capital	54,5	87,3	84,9	66,7
Vale do Paraíba e Litoral Norte	63,2	53,4	41,8	46,7
Sorocaba	85,9	94,9	95,5	92,1
Central	78,1	77,7	58,1	64,4
Mogiana	56,4	40,3	38,8	37,4
Paulista	65,8	58,8	50,2	58,6
Araraquarense	64,4	59,5	61,1	62,5
Noroeste	87,1	84,0	81,0	68,8
Alta Sorocabana	72,7	67,0	60,4	64,6
Baixa Sorocabana	75,5	83,7	86,7	92,4
Santos e Litoral Sul	3,4	16,1	1,2	18,2
Estado de São Paulo	65,9	62,4	55,6	61,5

Notas: Especialização da produção agrícola medida pelo índice definido no texto (ver equação na página 396). Cálculo baseado nos dados individuais de 36.266 propriedades rurais com área cultivada com milho no estado de São Paulo. Tamanhos das propriedades de acordo com a classificação de Prado Jr. (1945a) e Milliet (1939/1982a): i) pequenas propriedades: de 0 a 25 alqueires; ii) médias propriedades: de 25 a 100 alqueires; iii) grandes propriedades: de 101 a 500 alqueires, e iv) latifúndios: maiores que 500 alqueires. 1 alqueire = 2,42 hectares ou 24.200 metros quadrados.
Fonte: São Paulo (1906-1907).

¹⁶Exemplos deste tipo de comparação para a América Latina: Yarrington (1997, p.64); Solbrig (2008, p.508). Para a visão clássica da grande propriedade como domínio exclusivo do café no Brasil, Prado Jr. (1945/1969, pp.163-4 e 248-52); Costa (1966/1998, pp.75 e 522).

¹⁷Ver, por exemplo, Camargo (1952, p.191); Cano (1981, pp.60–61); Font (1990, pp.18–21); Castro (1987, pp.22–61).

Como é possível ver na tabela, as pequenas propriedades apresentavam um número relativo (65,9%) de unidades especializadas em milho superior ao das demais classificações de propriedades por tamanho no estado de São Paulo — tanto as médias (62,4%) e grandes propriedades (55,6%) como os latifúndios (61,5%). Esses dados demonstram que as pequenas propriedades possuíam uma tendência maior de especializarem-se em produtos voltados ao mercado interno, bastando comparar com o percentual visto anteriormente de pequenas propriedades (24,7%) especializadas em café. Os menores índices de especialização das pequenas propriedades — em relação à média do estado — foram registrados nas zonas tipicamente cafeeiras (Mogiana, Paulista e Araraquarense), Capital, Vale do Paraíba e Santos e Litoral Sul, essa última tradicionalmente especializada em arroz e produtos diversos.¹⁸ Ainda assim, os números acima também indicam que um elevado percentual de grandes propriedades especializava-se na produção de alimentos — 55,6% das grandes propriedades e 61,5% dos latifúndios. No caso das regiões do café, a proporção de latifúndios e grandes propriedades especializados em milho variou entre 37 e 63% do total. Nas regiões de Sorocaba e Baixa Sorocabana os percentuais foram ainda mais elevados, com cerca de 87% a 96% dos latifúndios e grandes propriedades especializando-se na lavoura de milho (Tabela 12).

Estas evidências complementam o que foi visto anteriormente a respeito do café, demonstrando que generalizações comuns sobre a posição de pequenas e grandes propriedades não são consistentes com o observado na economia agrária de São Paulo no início do século XX. Embora um maior número relativo de pequenas propriedades tenha se especializado na produção de mantimentos, sobretudo nas zonas com menor presença de café, a parcela de grandes propriedades e latifúndios especializada em milho e outros alimentos foi também substancial, provavelmente contribuindo para o atendimento ao consumo doméstico desses bens, conforme será visto a seguir.

6. PEQUENAS E GRANDES PROPRIEDADES NA OFERTA AGRÍCOLA

Embora um grande número de pequenas propriedades tenha se especializado em café, e um número ainda maior em alimentos, uma questão diferente é o peso relativo de sua produção no total da oferta dos bens agrícolas. A produção de café por parte das pequenas propriedades possuía alguma expressão quantitativa, como parecem sugerir os indicadores de área cultivada e especialização produtiva vistos anteriormente? Da mesma forma, qual era o peso da produção de café e alimentos nas grandes propriedades? A produção de alimentos era dominada por pequenas propriedades, como dá a entender boa parte da historiografia?

Conforme mostra a Tabela 13, no início do século XX a produção de café concentrava-se nas três zonas especializadas nesse produto: Mogiana (33,8%), Paulista (21,2%) e Araraquarense (15,8%). A antiga região Central (12,1%) e a nova área de expansão na Alta Sorocabana (10,4%) também exibiam uma participação expressiva na produção cafeeira. A zona pioneira do café no início do século XIX, constituída pelo Vale do Paraíba e Litoral Norte, já havia declinado para uma produção marginal (5,4%), enquanto as regiões restantes tinham participação negligível.

A distribuição da produção do café segundo o tamanho das propriedades agrícolas é resumida na Tabela 14. As grandes propriedades (46,6%) concentravam a maior parte da produção cafeeira em São Paulo, seguidas por médias propriedades (23,5%) e latifúndios (22,3%), ficando as pequenas unidades em posição marginal (7,6%) no total da produção do estado. Considerando as zonas cafeeiras principais, as grandes propriedades e os latifúndios somados produziam 59% do total do café na Araraquarense, 72% na Mogiana e 77% na Paulista. As médias propriedades respondiam por uma parcela substantiva do café sobretudo na Araraquarense (com 32,6% do total), superando inclusive os latifúndios (15,0%). As propriedades médias também produziam mais café do que os latifúndios nas regiões produtoras secundárias: a Central (24,8% contra 11,1%) e Alta Sorocabana (24,3% contra 17,7%), além do Vale do Paraíba e Litoral Norte (30,7% contra 12,8%).

¹⁸Os dados desses produtos relativos a Santos e Litoral Sul não são apresentados no texto.

**Tabela 13.** Produção de Café por Regiões, São Paulo, 1904-1905 (em %).

Regiões	Produção de Café
Capital	0,03
Vale do Paraíba e Litoral Norte	5,4
Sorocaba	0,69
Central	12,1
Mogiana	33,8
Paulista	21,2
Araraquarense	15,8
Noroeste	0,29
Alta Sorocabana	10,4
Baixa Sorocabana	0,31
Santos e Litoral Sul	0,08
Estado de São Paulo	100

Fonte: São Paulo (1906-1907).

Tabela 14. Produção de Café por Classes de Tamanho das Propriedades, São Paulo, 1904-1905 (em %).

Regiões	Pequena Propriedade	Média Propriedade	Grande Propriedade	Latifúndio	Total
Capital	63,9	26,5	9,6	0,0	100,0
Vale do Paraíba e Litoral Norte	16,3	30,7	40,1	12,8	100,0
Sorocaba	11,9	30,2	41,5	16,5	100,0
Central	9,1	24,8	55,1	11,1	100,0
Mogiana	7,2	20,8	44,5	27,6	100,0
Paulista	4,9	18,0	47,3	29,7	100,0
Araraquarense	8,6	32,6	43,8	15,0	100,0
Noroeste	3,6	11,8	22,5	62,1	100,0
Alta Sorocabana	6,8	24,3	51,3	17,7	100,0
Baixa Sorocabana	13,3	29,3	35,1	22,3	100,0
Santos e Litoral Sul	2,3	22,9	18,5	56,3	100,0
Estado de São Paulo	7,6	23,5	46,6	22,3	100,0

Notas: Cálculo baseado nos dados individuais de 21.010 propriedades rurais com produção registrada de café no estado de São Paulo. Tamanhos das propriedades de acordo com a classificação de Prado Jr. (1945a) e Milliet (1939/1982a): i) pequenas propriedades: de 0 a 25 alqueires; ii) médias propriedades: de 25 a 100 alqueires; iii) grandes propriedades: de 101 e 500 alqueires, e iv) latifúndios: maiores que 500 alqueires. 1 alqueire = 2,42 hectares ou 24.200 metros quadrados.

Fonte: São Paulo (1906-1907).

Uma explicação para os percentuais relativamente altos dos latifúndios na produção cafeeira da Mogiana e Paulista pode ter sido o fato dessas regiões estarem, na época, entre as zonas mais novas e dinâmicas na produção do café. Sérgio Milliet (1939/1982b), por exemplo, argumentou que o declínio do café nas regiões mais antigas (Vale do Paraíba e Central) levou à diversificação agrícola e à repartição de terras, aumentando a presença das pequenas e médias lavouras. Contudo, diante da ausência de dados comparáveis de períodos anteriores, não é possível afirmar se a proeminência do latifúndio frente à média propriedade na Mogiana e na Paulista deveu-se ao dinamismo da lavoura cafeeira nessas regiões de expansão recente ou a características institucionais locais de posse da terra.¹⁹ Além disso, as médias propriedades na Alta Sorocabana, por exemplo, produziam relativamente mais café (24,3%) do que os latifúndios (17,7%), mesmo sendo essa região, na época, uma área de fronteira da cafeicultura (Tabela 14).

No caso de um produto típico destinado ao mercado doméstico, o milho, a produção revela-se menos concentrada e com uma distribuição distinta daquela do café entre as regiões do estado de São Paulo, de acordo com a Tabela 15.²⁰ A antiga zona Central aparece como a maior região produtora (19,1%), acompanhada pelas regiões cafeeiras da Paulista (16,9%), Mogiana (15,6%) e Araraquarense (14,0%). Novamente contrastando com o cultivo de café, outras regiões não cafeeiras registraram produção expressiva: Vale do Paraíba (10,4%), Baixa Sorocabana (7,4%) e Sorocaba (6,5%), além da Alta Sorocabana (8,6%).

Tal como ocorria com o café, contudo, a maior produção de milho teve lugar nas grandes propriedades, com 33,9% da colheita no estado de São Paulo (Tabela 16). As médias propriedades também vinham em segundo lugar com 27,8%, seguidas pelos latifúndios, que chegavam a 20,5% da produção estadual. As pequenas propriedades apresentaram a menor contribuição relativa (17,8%) na oferta de milho, mas sua participação era, ainda assim, mais de duas vezes superior à observada no caso do café (17,8% contra 7,6%).²¹ O cultivo do milho era uma atividade majoritariamente de pequenas propriedades apenas em

Tabela 15. Produção de Milho por Regiões, São Paulo, 1904-1905 (em %).

Regiões	Produção de Milho
Capital	0,99
Vale do Paraíba e Litoral Norte	10,4
Sorocaba	6,5
Central	19,1
Mogiana	15,6
Paulista	16,9
Araraquarense	14
Noroeste	0,59
Alta Sorocabana	8,6
Baixa Sorocabana	7,4
Santos e Litoral Sul	0,06
Estado de São Paulo	100

Fonte: São Paulo (1906-1907).

¹⁹Para o efeito de características institucionais locais sobre a concentração da propriedade em parte da região da Mogiana, ver Bacellar (1999).

²⁰Os resultados descritos a seguir para a produção de milho são semelhantes aos do feijão, cujos dados não são apresentados aqui por razões de espaço.

²¹Esses resultados conflitam com o que Font (1995) afirma que seria a estrutura agrária de São Paulo alguns anos depois: segundo o autor, os pequenos proprietários rurais dominavam a produção para o mercado interno em São Paulo na década de 1920. Ver também Font (1990, cap.1).

**Tabela 16.** Produção de Milho por Classes de Tamanho das Propriedades, São Paulo, 1904-1905 (em %).

Regiões	Pequena Propriedade	Média Propriedade	Grande Propriedade	Latifúndio	Total
Capital	46,9	31,0	19,1	3,0	100,0
Vale do Paraíba e Litoral Norte	33,8	29,3	27,4	9,6	100,0
Sorocaba	10,4	33,2	38,0	18,2	100,0
Central	19,6	30,7	32,7	17,1	100,0
Mogiana	16,1	24,4	38,8	20,7	100,0
Paulista	11,7	21,1	38,5	28,6	100,0
Araraquarense	20,1	34,0	32,8	13,1	100,0
Noroeste	13,0	33,6	23,6	29,7	100,0
Alta Sorocabana	17,2	30,0	34,4	18,5	100,0
Baixa Sorocabana	8,3	20,1	26,1	45,5	100,0
Santos e Litoral Sul	9,2	52,2	25,7	13,0	100,0
Estado de São Paulo	17,8	27,8	33,9	20,5	100,0

Notas: Cálculo baseado nos dados individuais de 37.163 propriedades rurais com produção registrada de milho no estado de São Paulo. Tamanhos das propriedades de acordo com a classificação de Prado Jr. (1945a) e Milliet (1939/1982a): *i*) pequenas propriedades: de 0 a 25 alqueires; *ii*) médias propriedades: de 25 a 100 alqueires; *iii*) grandes propriedades: de 101 e 500 alqueires, e *iv*) latifúndios: maiores que 500 alqueires. 1 alqueire = 2,42 hectares ou 24.200 metros quadrados.

Fonte: São Paulo (1906-1907).

duas regiões, a Capital e o Vale do Paraíba e Litoral Norte, onde alcançavam 46,9% e 33,8% do total produzido, respectivamente. Em outras duas zonas, Central e Araraquarense, as pequenas propriedades produziam mais do que os latifúndios, algo que não ocorria nas demais regiões. Já na Baixa Sorocabana a maior produção relativa era encontrada nos latifúndios, uma característica que não se observa em qualquer outra região de São Paulo na época. Esses resultados díspares tornam difícil generalizar as características da produção de milho (e feijão) de uma região para o conjunto do estado de São Paulo.

De uma maneira geral, portanto, a produção de um bem tipicamente destinado ao mercado doméstico como o milho apresentava uma distribuição regional distinta da produção cafeeira, mas não se distinguia de forma apreciável da lavoura do café no que diz respeito à contribuição das grandes propriedades à oferta agrícola.

7. CONCLUSÕES

Duas conclusões gerais podem ser extraídas da análise realizada nas seções anteriores. Essas conclusões corroboram os estudos de geógrafos e historiadores das primeiras décadas do século XX que retrataram um quadro diversificado de especializações agrícolas no estado de São Paulo em meio à grande expansão cafeeira, mas ao mesmo tempo adicionam novas evidências e dimensões que complementam tais estudos. Em primeiro lugar, a especialização exclusiva no cultivo do café no início do século XX restringiu-se àquela que era então a maior região exportadora do mundo — a Mogiana. Outras zonas cafeeiras especializadas no café (Araraquarense e Paulista) possuíam estruturas agrárias mais diversificadas, o que determinou o fato aparentemente paradoxal de que as três regiões especializadas em café estavam entre aquelas com a maior área cultivada com os principais produtos (arroz, milho e feijão) destinados ao mercado doméstico.

As demais regiões especializaram-se em produtos do mercado interno, mesmo que o café cultivado em zonas como a Central, o Vale do Paraíba e a Alta Sorocabana tenha ocupado uma parcela significativa da área plantada do produto no conjunto do estado. Deste modo, uma análise mais sensível às variações regionais possibilita identificar características importantes da estrutura e dos mercados agrários que em geral são suprimidas por análises agregadas da economia de São Paulo durante a expansão cafeeira.

A segunda conclusão geral diz respeito às propriedades rurais individuais. As grandes propriedades foram aquelas por excelência especializadas na produção do café, mas as pequenas propriedades também apresentaram altos índices de especialização no produto, particularmente nas zonas cafeeiras, indicando que a produção de exportação era atraente para proprietários de todos os tamanhos. Por consequência, a noção de que o cultivo do café era tipicamente o domínio das grandes propriedades, com elevadas economias de escala e capitalização, deve ser pelo menos qualificada tanto nas comparações tradicionais do café na América Latina (Brasil em contraste com Venezuela, Costa Rica e partes da Colômbia) como nas análises específicas sobre o Brasil. De forma análoga, a despeito da tendência de maior especialização das pequenas propriedades nos produtos de consumo doméstico, as grandes propriedades apresentaram altos índices de especialização no cultivo de produtos típicos do mercado interno, como arroz, feijão e milho.

As informações relativas ao volume de produção agrícola reforçam essa última conclusão. Os dados de produção mostram que as grandes propriedades ocupavam um lugar central tanto no cultivo do produto de exportação quanto em alguns dos mais importantes produtos do mercado interno no estado de São Paulo do início do século XX. Os latifúndios também detinham uma participação significativa da oferta de ambos tipos de produtos, embora estivessem em uma posição secundária em relação às médias propriedades. Na época, a oferta desses produtos para o mercado interno só foi dominada por pequenas propriedades em regiões específicas, a Capital e o Vale do Paraíba e Litoral Norte. Em outras, a produção de mantimentos podia ser majoritariamente realizada em grandes propriedades ou até latifúndios, como na Baixa Sorocabana.

Estes resultados demonstram que a generalização para o estado de São Paulo ou para o Brasil de evidências encontradas em localidades ou regiões específicas pode levar a uma imagem distorcida da posição ocupada por pequenas e grandes propriedades do século XIX e início do século XX. E que é preciso revisar noções arraigadas — inclusive em trabalhos recentes — sobre a natureza das estruturas agrárias dos produtos de exportação e do mercado interno na história brasileira.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bacellar, C. d. A. P. (1999). Uma rede fundiária em transição. In C. de Almeida Prado Bacellar & L. R. Brioschi (Eds.), *Na estrada do Anhanguera: Uma visão regional da história paulista*. São Paulo: Humanitas/FFLCH-USP.
- Bassanezi, M. S. C. B., & Francisco, P. M. S. B. (Eds.). (2003). *Estado de São Paulo: Estatística agrícola e zootécnica, 1904–1905* [CD-ROM]. Campinas: Nepo/Unicamp.
- Brasil, Directoria Geral de Estatística. (1876). *Recenseamento geral do Império de 1872*. Rio de Janeiro: Typ. Leuzinger.
- Brasil, Directoria Geral de Estatística. (1917). *Anuario estatístico do Brazil, 1908–1912* (Vol. II). Rio de Janeiro: Typographia da Estatística.
- Brasil, Directoria Geral de Estatística (1926). *Recenseamento do Brasil* (Realizado em 1 de Setembro de 1920). Vol. IV, 1ª Parte. Rio de Janeiro: Typ. da Estatística.
- Camargo, J. F. d. (1952). *Crescimento da população no estado de São Paulo e seus aspectos econômicos* (Vol. 1). São Paulo: Universidade de São Paulo.
- Cano, W. (1981). *Raízes da concentração industrial em São Paulo* (2ª ed.). T.A. Queiroz.
- Castro, H. M. M. d. (1987). *Ao sul da história*. São Paulo: Brasiliense.
- Cobra, A. N. (1923). *Em um recanto do sertão paulista*. São Paulo: Typ. Hennies Irmãos.
- Colistete, R. P., & Lamounier, M. L. (2014). *Land inequality in a coffee economy: São Paulo during the early twentieth century* (Working Papers Series Nº 2014-1). São Paulo: Department of Economics – FEA/USP. Disponível em: http://www.fea.usp.br/feaecon/RePEc/documentos/Colistete_Lamounier01WP.pdf
- Costa, E. V. d. (1998). *Da senzala à Colônia* (4ª ed.). São Paulo: Editoria Unesp. (Obra original de 1966)
- Dean, W. (1997). *With broadax and firebrand: The destruction of the Brazilian atlantic coastal forest*. Los Angeles: University of California Press.



- de Castro, H. M. M. d. (1988). Beyond masters and slaves: Subsistence agriculture as a survival strategy in Brazil during the second half of the nineteenth century. *Hispanic American Historical Review*, 68, 461–89.
- Deffontaines, P. (1936a). Pays et paysages de l'État de Saint-Paul (Brésil). Première esquisse de division régionale. *Annales de Géographie*, 45, 50–71.
- Deffontaines, P. (1936b). Pays et paysages de l'État de Saint-Paul (Brésil). Première esquisse de division régionale. *Annales de Géographie*, 45, 160–174.
- Denis, P. (1909). *Brazil*. London: T. Fisher Unwin.
- Font, M. (1990). *Coffee, contention, and change in the making of modern Brazil*. Oxford: Blackwell.
- Font, M. (1995). Labor system and collective action in a coffee export sector: São Paulo. In W. Roseberry, L. Gudmundson & M. S. Kutschbach (Eds.), *Coffee, society, and power in Latin America*. Baltimore: Johns Hopkins University Press.
- Gudmundson, L. (1989). Peasant, farmer, proletarian: Class formation in a smallholder coffee economy, 1850–1950. *Hispanic American Historical Review*, 69, 221–57.
- Holanda, S. B. d. (1994). Uma civilização do milho. In *Caminhos e fronteiras* (3ª ed.). São Paulo: Companhia das Letras. (Obra original de 1956)
- Holloway, T. (1974). *Migration and mobility: Immigrants as laborers and landowners in the coffee zone of São Paulo, 1886-1934* (PhD diss.). University of Wisconsin, Madison, WI.
- James, P. (1932). The coffee lands of Southeastern Brazil. *Geographical Review*, 22, 225–244.
- Love, J. (1982). *A locomotiva: São Paulo na federação brasileira, 1889–1937*. Rio de Janeiro: Paz e Terra.
- Milliet, S. (1982a). O desenvolvimento da pequena propriedade no estado de São Paulo. In *Roteiro do café e outros ensaios: Contribuição para o estudo da história econômica e social do Brasil* (4ª ed.). Hucitec. (Obra original de 1939)
- Milliet, S. (1982b). Roteiro do café. In *Roteiro do café e outros ensaios: Contribuição para o estudo da história econômica e social do Brasil* (4ª ed.). Hucitec. (Obra original de 1939)
- Monbeig, P. (1937). Les zones pionnières de l'État de São Paulo. *Annales d'Histoire Économique et Sociale*, 46, 343–65.
- Monbeig, P. (1945-1946). A divisão regional do estado de São Paulo. *Anais da Associação dos Geógrafos Brasileiros*, 1, 19–36.
- Monbeig, P. (1957). As estruturas agrárias da faixa pioneira paulista. In *Novos estudos de geografia humana brasileira*. São Paulo: Difusão Européia do Livro.
- Monbeig, P. (1998). *Pioneiros e fazendeiros de São Paulo* (2ª ed.) [Tradução da tese original de 1949]. São Paulo: Hucitec.
- Nakamura, R., & Paul, C. M. (2009). Measuring agglomeration. In R. Capello & P. Nijkamp (Eds.), *Handbook of regional growth and development theories*. Cheltenham: Edward Elgar.
- Pestana, P. (1923). *A expansão econômica do estado de S. Paulo num século, 1822–1922*. São Paulo: Secretaria da Agricultura, Comércio e Obras Públicas do Estado de São Paulo.
- Platt, R. (1935). Coffee plantations of Brazil: A comparison of occupance patterns in established and frontier areas. *Geographical Review*, 25, 231–39.
- Prado Jr., C. (1945a). Distribuição da propriedade fundiária rural no estado de São Paulo [Publicado originalmente em *Geografia*, 1, 1935]. *Boletim Geográfico*, 29, 692–700.
- Prado Jr., C. (1945b). O fator geográfico na formação e no desenvolvimento da cidade de São Paulo [Publicado originalmente em *Geografia*, 3, 1935]. *Boletim Geográfico*, 31, 920–31.
- Prado Jr., C. (1969). *História econômica do Brasil* (11ª ed.). São Paulo: Brasiliense. (Obra original de 1945)
- Roseberry, W. (1995). Introduction. In W. Roseberry, L. Gudmundson & M. S. Kutschbach (Eds.), *Coffee, society, and power in latin america*. Baltimore: Johns Hopkins University Press.
- Sampaio, T. (1890). Considerações geográficas e econômicas sobre o Valle do Rio Paranapanema. *Boletim da Comissão Geographica e Geologica do Estado de S. Paulo*, 4.
- Sampaio, T. (1900–1901). S. Paulo no seculo XIX. *Revista do Instituto Historico e Geographico de São Paulo*, VI, 159-205.
- Singh, J., & Dhillon, S. (1984). *Agricultural geography*. New Dheli: Tata McGraw-Hill.

- Solbrig, O. (2008). Structure, performance, and policy in agriculture. In *The cambridge economic history of Latin America* (Vol. II). Cambridge: Cambridge University Press.
- São Paulo, Comissão Central de Estatística. (1888). *A província de São Paulo em 1888: Relatório apresentado ao Exm. Sr. Presidente da Província de São Paulo*. Typographia King.
- São Paulo, Secretaria de Agricultura, Commercio e Obras Publicas do Estado de São Paulo. (1906-1907). *Estatística agrícola e zootechnica do estado de São Paulo no anno agrícola de 1904–1905*. São Paulo: Tip. Brazil.
- Topik, S., & Clarence-Smith, W. G. (2003). Conclusion: New propositions and a research agenda. In W. G. Clarence-Smith & S. Topik (Eds.), *The global coffee economy in Africa, Asia, and Latin America, 1500–1989*. Cambridge, UK: Cambridge University Press
- Van Delden Laerne, C. F. (1885). *Rapport sur la culture du café en Amérique, Asie et Afrique*. La Haye: Martinus Hijhoff.
- Yarrington, D. (1997). *A coffee frontier: Land, society, and politics in Duaca, Venezuela, 1830–1936*. Pittsburgh: University of Pittsburgh Press.



APÊNDICE. CLASSIFICAÇÃO DAS REGIÕES DE SÃO PAULO E SEUS MUNICÍPIOS

Zona	Municípios
Capital	Capital (1558), Conceição dos Guarulhos (Guarulhos, 1880), Cotia (1856), Itapeverica (1877), Juqueri (Mairiporã, 1889), Santo Amaro (1832), São Bernardo (São Bernardo do Campo, 1889)
Vale do Paraíba e Litoral Norte	Areias (1816), Bananal (1832), Santo Antônio da Bocaina (Cachoeira Paulista, 1880), Buquira (Monteiro Lobato, 1880), Caçapava (1855), Cruzeiro (1871), Cunha (1785), Guararema (1898), Guarantiguetá (1651), Jacareí (1656), Jambeiro (1876), Jataí (1887), Lagoinha (1880), Lorena (1788), Mogi das Cruzes (1611), Natividade (Natividade da Serra, 1863), Paraibuna (1832), Patrocínio de Santa Izabel (Igaratá, 1873), Pindamonhangaba (1705), Pinheiros (Lavrinhas, 1881), Queluz (1842), Redenção (Redenção da Serra, 1877), São José do Barreiro (1859), São José dos Campos (1767), Salesópolis (1857), Santa Branca (1856), Santa Isabel (1832), São Bento do Sapucaí (1858), São Luis do Paraitinga (1773), Silveiras (1864), Taubaté (1645), Tremembé (1896), Vila Vieira do Piquete (Piquete, 1891), Caraguatatuba (1857), São Sebastião (1636), Ubatuba (1637), Vila Bela (Ilhabela, 1806)
Sorocaba	Araçatiguama (1874), Campo Largo de Sorocaba (Araçoiaba da Serra, 1857), Espírito Santo da Boa Vista (Angatuba, 1885), Guareí (1880), Itapetininga (1771), Parnaíba (Santana de Parnaíba, 1625), Piedade (1857), Pilar (Pilar do Sul, 1891), São Roque (1832), Sarapuí (1872), Sorocaba (1661), Tatui (1844), Una (Ibiúna, 1857)
Central	Atibaia (1769), Bragança (Bragança Paulista, 1797), Cabreúva (1859), Campinas (1797), Capivari (1832), Indaiatuba (1859), Itatiba (1857), Itú (1654), Jundiaí (1655), Monte Mor (1871), Nazaré (Nazaré Paulista, 1850), Piracaia (1859), Piracicaba (1821), Porto Feliz (1797), Rio das Pedras (1894), Salto de Itú (Salto, 1889), Santa Bárbara (Santa Bárbara d'Oeste, 1869), São João do Curralinho (Joanópolis, 1895), São Pedro (1881)
Mogiana	Amparo (1857), Batatais (1839), Caconde (1864), Cajurú (1865), Casa Branca (1841), Cravinhos (1897), Espírito Santo do Pinhal (1877), Franca (1821), Itapira (1858), Ituverava (1885), Jardinópolis (1898), Mococa (1871), Mogi-Guaçu (1877), Mogi-Mirim (1769), Nuporanga (1885), Patrocínio do Sapucaí (Patrocínio Paulista, 1885), Pedreiras (Pedreira, 1896), Ribeirão Preto (1871), Santa Rita do Paraíso (Igarapava, 1873), Santo Antônio da Alegria (1885), São João da Boa Vista (1859), São José do Rio Pardo (1885), São Simão (1865), Serra Negra (1859), Sertãozinho (1896), Socorro (1871), Tambaú (1898)
Paulista	Anápolis (Analândia, 1897), Araraquara (1832), Araras (1871), Barretos (1885), Bebedouro (1894), Belém do Descalvado (Descalvado, 1865), Jaboticabal (1867), Leme (1895), Limeira (1842), Pirassununga (1865), Pitangueiras (1893), Porto Ferreira (1896), Rio Claro (1845), Santa Cruz da Conceição (1898), Santa Cruz das Palmeiras (1885), Santa Rita do Passa Quatro (1885), São Carlos do Pinhal (1865)
Araraquarense	Bariri (1890), Boa Esperança (1898), Boa Vista das Pedras (Itápolis, 1891), Brotas (1859), Dois Córregos (1874), Dourados (Dourado, 1897), Ibitinga (1890), Jaú (1866), Matão (1898), Mineiros (Mineiros do Tietê, 1898), Monte Alto (1895), Pederneiras (1891), Ribeirão Bonito (1890), Ribeirãozinho (Taquaritinga, 1892), São João da Bocaina (Bocaina, 1891), São José do Rio Preto (1894)
Noroeste	Bauru (1887)
Alta Sorocabana	Avaré (1875), Botucatu (1855), Campos Novos do Paranapanema (Campos Novos Paulista, 1885), Espírito Santo do Turvo (1885), Fartura (1891), Lençóis (Lençóis Paulista, 1865), Pereiras (1889), Piraju (1880), Remédios do Tietê (Anhembí, 1891), Rio Bonito (Bofete, 1880), Santa Bárbara do Rio Pardo (Águas de Santa Bárbara, 1876), Santa Cruz do Rio Pardo (1876), São João de Itatinga (Itatinga, 1896), São Manoel do Paraíso (São Manuel, 1885), São Paulo dos Agudos (Agudos, 1898), São Pedro do Turvo (1891), Tietê (1842)
Baixa Sorocabana	Apiáí (1771), Bom Sucesso (1885), Capão Bonito do Paranapanema (Capão Bonito, 1857 e 1868), Faxina (Itapeva, 1769), Iporanga (1873 e 1936), Itaporanga (1871), Itararé (1896), Lavrinhas da Faxina (Itaberá, 1891), Ribeirão Branco (1892), São Miguel Arcanjo (1889), Santo Antonio da Boa Vista (Itaí, 1891), Xiririca (Eldorado, 1842)
Santos e Litoral Sul	Cananéia (1587), Conceição de Itanhaém (Itanhaém, 1549 ou 1579), Iguape (1635 ou 1638), Santos (1545), São Vicente (1700)

Nota: os anos de fundação referem-se às datas em que foram publicadas as leis de criação das vilas, que não necessariamente coincidem com as suas datas de instalação.

Fontes: Camargo (1952); Holloway (1974); Milliet (1939/1982b). Informações sobre datas de fundação de acordo com São Paulo (1888) e Seade, "Histórico da Formação dos Municípios de São Paulo" (<http://www.seade.gov.br/produtos/500anos>).

The Naive Central Banker

Marcelo de Carvalho Griebeler*

Contents: 1. Introduction; 2. The Benchmark Case: The Stabilizer Central Banker; 3. The naive central banker; 4. Optimal monetary policy; 5. The gain of deviation under commitment; 6. Concluding Remarks; Appendix A. The stabilizer central banker's optimal monetary policy; Appendix B. The naive central banker's optimal monetary policy; Appendix C. Omitted proofs; Appendix D. Graphics.

Keywords: Naive Central Banker, Inflationary Bias, Monetary Policy.

JEL Code: E58, E52.

There has been in some countries a trend of assigning other functions to central banks besides price stability. The most suggested function to be added to monetary authority's obligations is to pursue economic growth or full employment. In this paper we characterize the behavior and analyse the optimal monetary policy of, what we call, a naive central banker. We describe the naive behavior as one that does face the inflation-unemployment trade-off, but it tries to minimize both variables simultaneously. Our findings, both under discretion and commitment, indicate that the naive central banker delivers lower expected inflation and inflation variance than the benchmark behavior whenever the economy is rigid enough. However, the degree of conservativeness also affects this result, such that the less conservative the naive policymaker, the more rigidity is necessary.

Existe, em alguns países, uma tendência de atribuir outras funções ao banco central além da estabilização de preços. A função mais sugerida a ser incorporada nas obrigações da autoridade monetária é a perseguição do crescimento econômico ou do pleno emprego. Neste artigo caracterizamos o comportamento e analisamos a política monetária ótima do chamado banqueiro central ingênuo. Descrevemos o comportamento ingênuo como aquele que enfrenta o trade-off inflação-desemprego, mas tenta minimizar ambas variáveis simultaneamente. Nossos resultados, tanto sob discrição quanto sob commitment, indicam que o banqueiro central ingênuo gera inflação esperada e variância inflacionária menores do que o comportamento de referência, sempre que a economia apresentar rigidez suficiente. Entretanto, o grau de conservadorismo também afeta este resultado, tal que quanto menos conservador for o banqueiro ingênuo, mais rigidez é necessária.

* Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas. Av. João Pessoa, 52, sala 33, 3º andar, Centro, Porto Alegre-RS, Brasil. CEP 90040-000. E-mail: marcelo.griebeler@ufrgs.br



1. INTRODUCTION

There has been in some countries a trend of assigning other functions to central banks besides price stability. This is particularly true in countries where the central banks are not independent and thus are potential targets of political pressure. The most suggested function to be added to monetary authority's obligations is to pursue economic growth or full employment. The effects of the recent financial crisis, mainly the decreasing of GDP and the increasing of unemployment, may be considered a possible explanation for this suggestion. Based on the Keynesian idea of stimulating the economy by using fiscal and monetary policy in period of crisis, some policymakers in those countries have cogitated to use the central bank as an "agent of development".¹

The idea of central banks having to pursue economic growth is not new. While we can find the theoretical roots of such a controversy on the debate about the neutrality of money and whether monetary policy should be used to affect output, the history of central banking tells how practice has evolved over time.² Epstein (2005), for example, presents an historical account of central bank's action both in developed and developing countries. His conclusion is that the dominant "best practice" approach to central banking is dramatically different from the historically dominant theory and practice. In fact, it is argued all central banks have engaged in some kind of industrial policy or "selective targeting" in order to promote growth, and the only difference lies in which sector they have incentivized.

It is worthy to note the above historical view is not consentaneous. Orphanides (2012) is one of the authors which claim that throughout the existence of central banks their main objective has remained the same: stability. It is argued that what has been evolved over time is not central bank's function or objectives but the central banker's understanding on how to achieve and maintain stability. Even in financial crisis periods, the activist approach to monetary policy—one which wants to guide the economy towards attainment of its ideal output level—has not been the one chosen by central bankers, according to the author. Although the referred stability covers price, financial and economic dimensions, the point made is clear, namely the central bank's behavior has changed too little over time.

However, the new trend of expanding central bank's function is not just about making monetary policy more or less active. Instead the idea is to assign more *statutory* functions or goals for the central bank. This implies, for example, that promoting full employment would be in the same level of priority than pursuing price stability. One can say that a central banker who tries to achieve both targets at the same time, although he is aware about the role of the inflation-unemployment trade-off expressed by the Phillips Curve, presents a naive behavior.

But what are the economic effects of the monetary policy of a central bank which must fight inflation and stimulate output simultaneously? Literature both empirical and theoretical have not studied optimal monetary policy under those circumstances. In fact, there is a plenty of researchers exploring policies of the so called stabilizer central banker (about the nomenclature, see Griebeler, 2015), whose preferences are represented by a loss function quadratic both in inflation and output (e.g. Backus & Driffill, 1985; Ball, 1995; Canzoneri, 1985; Chari & Kehoe, 1990; Cukierman & Liviatan, 1991), in line with the seminal studies by Barro & Gordon (1983a, 1983b). But, to the best of our knowledge, none has considered a monetary authority which takes price and output stability as complementary targets.

In this paper we characterize the behavior and analyse the optimal monetary policy of, what we call, a naive central banker. We describe the naive behavior as one that does face the inflation-unemployment trade-off, but it tries to minimize both variables simultaneously. Specifically, it is defined as one that considers price and output stabilization as complementary "goods". In particular, we are interested in investigating the effects on the inflationary bias under both discretion and commitment. In order to

¹Observe that the trend we have mentioned is different from the controversy about the adoption of NGDP targeting versus price level targeting. While we are referring to a increase in the central bank's functions—pursuing both targets, for instance—, this controversy is about choosing one of the two regimes.

²For more on central banking and its evolution throughout history, see Blinder (2008, 2010); Capie, Fischer, Goodhart, & Schnadt (1995); Goodhart (2011).

do so, we use the simple Barro–Gordon’s framework, in which the central bank must minimize its loss function subject to a Lucas’s supply curve and the rational expectations assumption. The novelty that we introduce in this context is a loss function which represents that kind of complementarity relationship. However, since our model is not dynamic, it is just a first step toward studying the effects of the naive behavior. In fact, reputation issues are out of our framework, such that we cannot study how central banker is penalized according to his choices over time.

Our main findings concern to the comparison between the results of the naive and the stabilizer central banker, our benchmark, who considers price and output stabilization as substitute goals. First, both behaviors present similarities: there is an inflationary bias, and this bias depend primarily on how conservative is the central banker. In fact, when his level of conservativeness is maximum, the bias vanishes. Second, when both policymakers act with discretion, for given level of conservativeness of both central bankers, a necessary and sufficient condition for the naive one delivers higher both expected inflation and inflation variance is that output be inelastic enough with relation to inflation surprise. Third, this result is also true under non-binding commitment, but now we have a difference, namely as the elasticity goes to infinity, inflation delivered by both central bankers converges. Finally, we compare the incentives to cheat the commitment that both face and found that, for the same 1% of deviation from the inflation target, the percentage gain in terms of welfare of the stabilizer central banker is twice as high as the gain of the naive one.

The above result indicates there is a case for the naive central banker. In fact, the naive policymaker delivers lower both expected inflation and inflation variance whenever the economy’s rigidity is high enough. By rigidity we mean mainly informational one, given we choose the Lucas supply curve, but we conjecture our results would change too little if we used a model with price or wage stickiness. Therefore, one could conclude that, in terms of inflation, it is preferable to have a naive individual conducting the monetary policy than a stabilizer one whenever such a rigidity is present in the economy. The drawback of such a conclusion is that how enough must be the economy’s rigidity depend on how conservative both central bankers are. Empirical evidence seems to indicate that a naive central banker would be a easier target of political pressure, what would reflect on a more populist behavior. Thus, the threshold of rigidity would be larger, decreasing the range in which there is a case for a naive central banker. In short, given the stabilizer central banker’s level of conservativeness, the less conservative the naive central banker, the more rigid must be the economy.

Our findings are related to all the literature about inflationary bias and the potential solutions for this problem. Besides the studies which followed the seminal model of Barro & Gordon (1983a, 1983b), many of them mentioned in one the above paragraphs, there are several slightly modified versions of the original one which are close to our analysis. Examples are works which consider persistence in output and asymmetric preferences for the central bank (e.g. Jonsson, 1997; Nobay & Peel, 2003; Ruge-Murcia, 2003). Another branch in this literature is the modern approach provided by the New Keynesian models (e.g. Clarida, Galí, & Gertler, 1999; Damjanovic, Damjanovic, & Nolan, 2008; Sauer, 2010). However, by analysing the behavior of a specific type of central banker, our results are more close to those of Guzzo & Velasco (1999); Jeger (2002); Lippi (2002), which found there is a case for a populist central banker, and Herrendorf & Lockwood (1997); Rogoff (1985), which focus on the behavior of the conservative central banker. A good survey on the inflationary bias problem may be found in Walsh (2010).

This paper is organized as following. In section 2 we present the benchmark behavior, namely the stabilizer central banker’s one. We analyse his optimal monetary policy under discretion and commitment, as well as the resulting inflationary bias and its determinants. The naive central banker is characterized in section 3. We also investigate the optimal choice under discretion and commitment in this section and in the next. Thus, in section 4, inflation and output delivered by both policymakers are compared. Section 5 presents a further comparison between the naive behavior and the benchmark by analysing the elasticity of their welfare with respect to deviations from the commitment of delivering the inflation target. We make final comments regarding the limitations of our model as well as to potential extensions in section 6. Details of the results’ derivation can be found in Appendix A, for the stabilizer



central banker, and Appendix B, for the naive one. Finally, Appendix C presents the propositions' proofs omitted throughout the text, and Appendix D shows some graphics of our findings.

2. THE BENCHMARK CASE: THE STABILIZER CENTRAL BANKER

We follow literature by assuming that central banker has preferences over only two variables, namely inflation, π , and output, y . Let Π and Y be the subsets of \mathbb{R} , then $(\pi, y) \in \Pi \times Y \subseteq \mathbb{R} \times \mathbb{R}$.³ Before we introduce the naive central banker, we define the benchmark case for future comparisons. Our choice is the so called stabilizer central banker, the most common behavior used in the literature. This kind of central banker wishes to minimize deviations of inflation from its target and of output from its social optimal level. These two goal are considered substitutes and the central banker gives different weights to each one in its loss function. Formally, we can use the definition suggested by Griebeler (2015):

Definition 1. Let λ be a parameter that measures the relative weight given to output y , y^* be the output target, π be the inflation, and π^* be the inflation target. We define preferences of a stabilizer central banker as $(\pi_1, y_1) \succeq_{STA} (\pi_2, y_2)$ if and only if $|\pi_1 - \pi^*| + \lambda |y_1 - y^*| \leq |\pi_2 - \pi^*| + \lambda |y_2 - y^*|$.

It is possible to show (Griebeler, 2015) that preferences like those defined above have an utility representation. Due to the non-smoothness of the absolute value function, it is useful to adopt the following well known functional form to represent them:

$$L_S(\pi, y) = \frac{1}{2} [(\pi - \pi^*)^2 + \lambda(y - y^*)^2].$$

As usual, throughout the paper, we assume policymaker's choices are constrained by a Lucas supply curve, given by

$$y = \bar{y} + a(\pi - \pi^e) + \epsilon, \quad a > 0, \tag{1}$$

where \bar{y} is the potential level of output (of long run), $\pi^e = \mathbb{E}[\pi]$ is the inflation expectation, ϵ is a supply shock, with mean and variance given by $\mathbb{E}[\epsilon] = 0$ and $\mathbb{V}[\epsilon] = \sigma_\epsilon^2$, respectively. The parameter a is a measure of output's sensitiveness with respect to inflation surprise. Assume we are considering variables in log, then a is the output's elasticity with respect to $(\pi - \pi^e)$. Furthermore, in order to have a source of inflationary bias in our model, we assume that $y^* > \bar{y}$, by the reasons suggested by Kydland & Prescott (1977), for example. With those assumption we are able to analyse optimal monetary policy.

2.1. Discretion

There is no uncertainty in our model, such that we can think of policymaker as choosing inflation directly. Thus, when central banker chooses inflation taking inflation expectation as given, his optimal choice results in

$$\mathbb{E}[\pi_S^D] = \pi^* + \lambda a(y^* - \bar{y}) \tag{2}$$

$$\mathbb{V}[\pi_S^D] = a^2 \lambda^2 \sigma_\epsilon^2. \tag{3}$$

Some conclusions may be drawn from (2) and (3). First, under discretion, there is an inflationary bias and its magnitude is given by $\mathbb{E}[\pi_S^D] - \pi^* = \lambda a(y^* - \bar{y})$. Therefore, the bias is determined by three factors: how high is the social optimal output in relation to the potential level, the degree of conservativeness of the policymaker, and the output's elasticity with respect to inflation surprise. Still, an ultra conservative central banker ($\lambda = 0$) is a sufficient condition to vanish the bias. Second, the

³Our results may be easily generalized to choice sets of any finite dimension. For example, they may apply to a central banker with preferences over inflation, output and exchange rate.

higher the value of a , the higher the volatility of inflation is. The same is true about parameter λ . Once again, an ultra conservative central banker yields inflation's variance equal to zero. Finally, by substituting π_S^D into (1), we find that $\mathbb{E}[y_S^D] = \bar{y}$ and $\mathbb{V}[y_S^D] = \sigma_\epsilon^2$, that is, there is no gain in terms of output by implementing π_S^D .

The value of the stabilizer central banker's expected loss under discretion is

$$\mathbb{E}[L_S(\pi_S^D, \bar{y} + \epsilon)] = \lambda(1 + \lambda a^2) \frac{[(y^* - \bar{y})^2 + \sigma_\epsilon^2]}{2}.$$

2.2. Commitment

Suppose now that the stabilizer central banker makes a binding commitment about what inflation will be before expected inflation is determined. Since the commitment is binding, $\pi^e = \pi$, which implies $y = \bar{y} + \epsilon$ and thus the policymaker's problem is to minimize $[(\pi - \pi^*)^2 + \lambda(\bar{y} + \epsilon - y^*)^2]/2$ by choosing π . The solution is simply $\mathbb{E}[\pi_S] = \pi^*$, with $\mathbb{V}[\pi_S] = 0$. However, the central banker could make an announcement of pursuing a determined level of inflation and then does not fulfill his promise. The reason for such a behavior is to create an inflation surprise and hence to increase output above its potential level.

Lets investigate what the stabilizer central banker gains by cheating the public. Suppose that his announcement makes $\pi^e = \pi^*$, such that now its optimal expected inflation and variance are given by

$$\mathbb{E}[\pi_S^C] = \pi^* + \frac{\lambda a}{1 + \lambda a^2} (y^* - \bar{y}) \quad (4)$$

$$\mathbb{V}[\pi_S^C] = \frac{\lambda^2 a^2}{(1 + \lambda a^2)^2} \sigma_\epsilon^2. \quad (5)$$

We can use the expressions above to find the expected output and its variance:

$$\mathbb{E}[y_S^C] = \frac{(\bar{y} + \lambda a^2 y^*)}{1 + \lambda a^2} \quad (6)$$

$$\mathbb{V}[y_S^C] = \frac{\sigma_\epsilon^2}{(1 + \lambda a^2)^2}.$$

The gain of cheating can be seen in terms of welfare,

$$\mathbb{E}[L_S(\pi_S^C, y_S^C)] = \frac{\lambda}{(1 + \lambda a^2)} \frac{[(y^* - \bar{y})^2 + \sigma_\epsilon^2]}{2},$$

because $\mathbb{E}[L_S(\pi_S^D, \bar{y} + \epsilon)] > \mathbb{E}[L_S(\pi_S^C, y_S^C)]$. So, there exists an incentive to the stabilizer central banker to create inflation surprise. In fact, it is possible to see that $\mathbb{E}[y_S^C] > \bar{y}$, which yields the improvement in his welfare. We will return to this topic later, when this incentive will be compared to one of the naive central banker. Also, observe the role of the degree of conservativeness of the stabilizer central banker in the output's behavior: the lower λ (more conservative) is, the lower the difference $\mathbb{E}[y_S^C] - \bar{y}$ and the higher the output's variance are.

3. THE NAIVE CENTRAL BANKER

We call a naive behavior one which considers stabilization of price and output complementary goals of policy. The reason for this name is that a central banker with such a behavior, even though knows the trade-off between the two goals, he tries to achieve them simultaneously. A naive central banker therefore prefers a state of the economy, defined by a pair (π, y) , to another if and only if both inflation and output are closer to their targets. Formally:



Definition 2. Let π^* and y^* be the inflation and the output desired by society, respectively. We define the preference of a naive central banker, \succsim_N , by: $(\pi_1, y_1) \succsim_N (\pi_2, y_2)$ if and only if $\max\{|\pi_1 - \pi^*|, \phi|y_1 - y^*|\} \leq \max\{|\pi_2 - \pi^*|, \phi|y_2 - y^*|\}$, where ϕ is a constant which indicates the degree of complementarity between price and output stabilizations.

Definition 2 clearly satisfies transitivity, completeness and continuity such that there exists a continuous utility function which represent such a preference. In fact, the behavior described above is similar to that of consumer choosing his bundle with two complementary goods. Thus, a function of Leontief's type may be a good choice to represent it.⁴ Observe, however, that, since there is a point of satiety, (π^*, y^*) , the indifference curves will be rectangles rather than the traditional "L" shape. Following the literature, we study the minimization problem of the central banker, such that \succsim_N may be represented by a loss (disutility) function:

$$L_N(\pi, y) = \max\{|\pi - \pi^*|, \phi|y - y^*|\}. \tag{7}$$

In some examples in the next sections we will assume $\phi = 1$, what we consider the most intuitive case and the nearest of the behavior described in the introduction, because the central banker gives the same weight to both targets. Nevertheless, our general form (7) allows different degrees of complementarity between inflation and output targets. In the rest of the paper we will say the higher the value of ϕ , the less conservative (the more populist) the naive central banker is. The reason for adopting such an informal definition is explained by an example. Consider two different naive central bankers, one with $\phi = 1$ and other with $\phi' = 2$. For the same pair (π, y) , the banker with $\phi' = 2$ always earns twice as much disutility than the one with $\phi = 1$. In fact, in the example they will achieve the same welfare only if output deviation of ϕ' is the half of the deviation of ϕ . In short, the higher the value of ϕ , the more weight is given to output stabilization by the naive central banker.

It is trivial to see that (7) does formally represent \succsim_N . An example is helpful to show intuitively that the naive behavior described in the introduction is well described by \succsim_N (and then by (7)). Suppose that $\pi^* = y^* = 2$, $\phi = 1$, and there are two available bundles, $(1, 1)$ and $(1, 1.5)$. By applying the definition of \succsim_N , we have $(1, 1) \sim_N (1, 1.5)$, which means that the naive central banker is indifferent between a pair of inflation and output in which both deviate 1 from their targets and a pair in which inflation deviates 1 and output deviates 0.5. This is exactly the essence of the behavior which "ignores" inflation-output trade-off: his welfare only increases when both aims approach their targets simultaneously, such that there is no opportunity of substitution between them. In contrast, $(1, 1.5) \succ_S (1, 1)$, because the stabilizer behavior does consider price and output stabilization as substitutes. In short, \succsim_N is a suitable way of capturing such a naive behavior. However, we can also show that the representation by (7) is not unique, such that $\hat{L}_N(\pi, y) = \max\{(\pi - \pi^*)^2, \phi(y - y^*)^2\}$ represents \succsim_N as well. Our choice by (7) is due to its straightforward interpretation, since absolute value is a measure of deviation. Moreover, the advantage in terms of differentiability presented by \hat{L}_N is needless, because the maximum function is already non-smooth.

Given that central banker's preferences are represented by a loss function rather than an utility one, his problem is a minimization:

$$\min_{\pi} L(\pi, y) \quad \text{subject to (1).}$$

To learn more about the behavior of the naive central banker, lets show that he faces an incentive to choose inflation higher than π^* in order to try to cause inflation surprise and then to affect positively the output. Suppose that people believe in the central banker's announcement that inflation target will be surely achieved, such that their inflation expectation is $\pi^e = \pi^*$. Further, suppose that the central banker does have discretion. Thus the Lucas supply curve (1) becomes $y = \bar{y} + a(\pi - \pi^*) + \epsilon$, such that the central banker's problem as function only of π is $L_N(\pi, y(\pi)) = \max\{|\pi - \pi^*|, \phi|\bar{y} - y^* + a(\pi - \pi^*) + \epsilon|\}$.

⁴More on preferences that can be representad by a Leontief utility function, see Voorneveld (2014).

The optimal inflation satisfies $|\pi - \pi^*| = \phi|\bar{y} - y^* + a(\pi - \pi^*) + \epsilon|$ and then we would have to consider two cases: $\pi - \pi^* = \phi[\bar{y} - y^* + a(\pi - \pi^*) + \epsilon]$ and $\pi - \pi^* = -\phi[\bar{y} - y^* + a(\pi - \pi^*) + \epsilon]$.

Figure 1 plots the rectangle of indifferences of the naive central banker and the Lucas supply curve. First, observe that the “smaller” is the rectangle—namely, its area—the lower is the loss, such that the minimum of $L(\pi, y)$ occurs in (π^*, y^*) . In addition, the degree of complementarity affects both height and width of the rectangles. Since ϕ multiplies the term of output stabilization in (7), for given π^* and y^* , the higher ϕ , the lower the height of the rectangles is. We can see that by observing that, for a fixed π_0 , the height of the rectangle is $h(\pi_0) = 2(\pi^* - \pi_0)/\phi$, such that $\lim_{\phi \rightarrow \infty} h(\pi_0) = 0$. On the other hand, the higher ϕ , the higher the width is: for a fixed y_0 , the width of the rectangle is $w(y_0) = 2\phi(y^* - y_0)$, such that $\lim_{\phi \rightarrow \infty} w(y_0) = \infty$ (assuming $y^* > y_0$).

By determining the perimeter of the rectangles of indifference, its height and width determine also how many different pairs of inflation and output yield the same welfare for the banker. A higher perimeter offers more options for a given level of disutility. Therefore, it is fundamental to know how conservative is the naive central banker in order to study his available choices. For example, for a central banker with $\phi = 2$, 1% of inflation deviation from its target yields the same loss that any value up to 0.5% of output deviation from its desirable level. Alternatively, if $\phi = \frac{1}{2}$, the loss yielded by 1% of deviation in inflation is the same of any deviation in output up to 2%.

Second, as the supply curve is upward-sloping, one can wonder whether the solution of the constrained problem may be any of four corners of the rectangles. The two cases we have to analyse are:

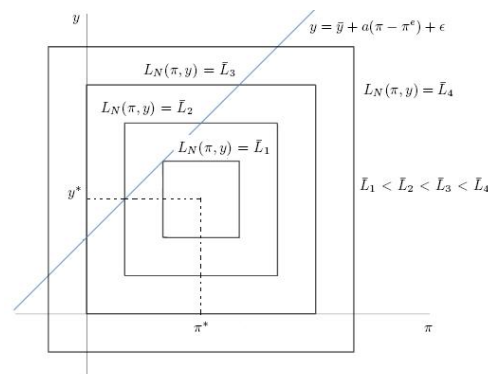
- (i) $\pi > \pi^*$ and $y > y^*$ or $\pi < \pi^*$ and $y < y^*$;
- (ii) $\pi > \pi^*$ and $y < y^*$ or $\pi < \pi^*$ and $y > y^*$.

In the next section we explore both cases and show which one will occur (see Proposition 2 below). For now, assume the case (ii) as true, such that $\pi - \pi^* = -\phi[\bar{y} - y^* + a(\pi - \pi^*) + \epsilon]$. Then, the optimal (expected) inflation and its variance are given by

$$\mathbb{E}[\pi_N^C] = \pi^* + \frac{\phi}{1 + a\phi}(y^* - \bar{y}) \tag{8}$$

$$\mathbb{V}[\pi_N^C] = \frac{\phi^2}{(1 + a\phi)^2} \sigma_\epsilon^2, \tag{9}$$

Figure 1. Rectangles of indifference of the naive central banker.





where one can see that there exists inflationary bias, because $y^* > \bar{y}$. This inflation surprise makes (expected) output be above its potential level:

$$\begin{aligned} \mathbb{E}[y_N^C] &= \bar{y} + \frac{a\phi}{1+a\phi}y^* \\ \mathbb{V}[y_N^C] &= (1-a\phi)^2\sigma_\epsilon^2. \end{aligned} \tag{10}$$

The expected central banker’s loss in this setting is then⁵

$$\begin{aligned} \mathbb{E}[L_N(\pi_N^C, y_N^C)] &= \mathbb{E}\left[\max\left\{\left|\frac{\phi}{1+a\phi}(y^* - \bar{y} - \epsilon)\right|, \phi\left|\frac{\bar{y} - y^* + \epsilon}{1+a\phi}\right|\right\}\right] \\ &= \frac{\phi}{1+a\phi}\mathbb{E}[|y^* - \bar{y} - \epsilon|]. \end{aligned}$$

The above results can be compared to those of the stabilizer central banker under the same conditions (non-binding commitment). For example, by comparing expected inflation and its variance, we find that $\mathbb{E}[\pi_N^C] \geq \mathbb{E}[\pi_S^C]$ and $\mathbb{V}[\pi_N^C] \geq \mathbb{V}[\pi_S^C]$ if and only if $\phi - \lambda a \geq 0$. This same condition also plays an important role when we compare the results related to output and its variability. Proposition 1 below brings more details about such a comparison.

Proposition 1. *Under non-binding commitment, the following statements are true:*

- (i) $\mathbb{E}[\pi_N^C] \geq \mathbb{E}[\pi_S^C]$ if and only if $a \leq \frac{\phi}{\lambda}$;
- (ii) $\mathbb{V}[\pi_N^C] \geq \mathbb{V}[\pi_S^C]$ if and only if $a \leq \frac{\phi}{\lambda}$;
- (iii) if $a \leq \frac{\phi}{\lambda}$, then $\mathbb{E}[y_N^C] \geq \mathbb{E}[y_S^C]$;
- (iv) as $a \rightarrow +\infty$, *ceteris paribus*, $\mathbb{E}[\pi_N^C] \rightarrow \mathbb{E}[\pi_S^C]$, $\mathbb{V}[\pi_N^C] \rightarrow \mathbb{V}[\pi_S^C]$, and $\mathbb{E}[y_N^C] > \mathbb{E}[y_S^C]$.

One way to interpret the results of Proposition 1 is by assuming that both levels of conservativeness, λ and ϕ , are fixed, such that relevant parameter for the analysis is the output’s elasticity with respect to inflation surprise, a . Thus, a necessary and sufficient condition for the naive central banker to deliver expected inflation and inflation variance higher than the stabilizer one is that the elasticity a be low enough. If, for instance, both central bankers have the same degree of conservativeness—even though their measures are not perfectly comparable—, such that $\phi/\lambda = 1$, then such a condition becomes: output has to be inelastic with relation to inflation surprise. Note also that the same condition is only sufficient for the expected output delivered by the naive central banker be higher than one delivered by the stabilizer one.

Another important point to note in the above results is that, despite the fact of high values of a implies both $\mathbb{E}[\pi_N^C] \leq \mathbb{E}[\pi_S^C]$ and $\mathbb{V}[\pi_N^C] \leq \mathbb{V}[\pi_S^C]$, as a increases the difference between the inflation results of both policymakers tends to zero. This means that, as the benefit of inflation surprise becomes greater, the behavior of both central bankers related to inflation converges. The same is not true about output: in many cases, for low enough values of a , the naive delivers higher output; for “intermediate” values, the stabilizer is who delivers higher output; and, as a increases (and exceeds a given value \bar{a}) the

⁵There is no closed form for $\mathbb{E}[|y^* - \bar{y} - \epsilon|]$. However, given that the absolute value is a convex function, we can apply the Jensen inequality and obtain a lower bound for $\mathbb{E}[L_N(\pi_N^C, y_N^C)]$. By doing so, we have $\mathbb{E}[|y^* - \bar{y} - \epsilon|] \geq |\mathbb{E}[y^* - \bar{y} - \epsilon]| = |y^* - \bar{y}| = y^* - \bar{y}$. Thus, $\mathbb{E}[L_N(\pi_N^C, y_N^C)] \geq \phi(y^* - \bar{y})/(1+a\phi)$.

naive again delivers higher output.⁶ One can visualize those results in figures D-1a, D-1b, D-1c, and D-1d in Appendix D, where are shown the differences $\mathbb{E}[\pi_N^C] - \mathbb{E}[\pi_S^C]$, $\mathbb{V}[\pi_N^C] - \mathbb{V}[\pi_S^C]$, $\mathbb{E}[y_N^C] - \mathbb{E}[y_S^C]$, and $\mathbb{V}[y_N^C] - \mathbb{V}[y_S^C]$, respectively, as a function of a . We also consider three possible cases: $\phi/\lambda = 1$, $\phi/\lambda > 1$ and $\phi/\lambda < 1$.

4. OPTIMAL MONETARY POLICY

4.1. Discretion

Under discretion, central banker chooses inflation taking expectations of inflation as given. Substituting (1) into (7), we have the unconstrained problem of minimization $L(\pi) = \max\{|\pi - \pi^*|, \phi|\bar{y} - y^* + a(\pi - \pi^e) + \epsilon|\}$. The solution is given by the points that satisfy $|\pi - \pi^*| = \phi|\bar{y} - y^* + a(\pi - \pi^e) + \epsilon|$. There are two possibilities to consider (cases (i) and (ii) in section 3). Proposition 2 states that the optimum always occurs either $\pi > \pi^*$ and $y < y^*$ or $\pi < \pi^*$ and $y > y^*$. Undoubtedly, the most common situation in practice is inflation higher than its target and output lower than its potential level, which henceforth we assume as true.

Proposition 2. *Consider the two possible cases of solution of $|\pi - \pi^*| = \phi|\bar{y} - y^* + a(\pi - \pi^e) + \epsilon|$, namely (i) $\pi - \pi^* = \phi|\bar{y} - y^* + a(\pi - \pi^e) + \epsilon|$ and (ii) $\pi - \pi^* = -\phi|\bar{y} - y^* + a(\pi - \pi^e) + \epsilon|$. Then, case (ii) always gives the minimum loss for the naive central banker.*

By using the result above, we may obtain the inflation as a function of public's expectation,

$$\pi = \pi^* + \frac{\phi a(\pi^e - \pi^*) + \phi(y^* - \bar{y}) - \phi\epsilon}{1 + \phi a}. \quad (11)$$

With rational expectations, people form their expectations by knowing central banker's incentive to stimulate inflation surprise, such that $\pi^e = \pi$. By substituting this into (11), we obtain the (expected) inflation of equilibrium and its variance:

$$\mathbb{E}[\pi_N^D] = \pi^* + \phi(y^* - \bar{y}) \quad (12)$$

$$\mathbb{V}[\pi_N^D] = \phi^2 \sigma_\epsilon^2. \quad (13)$$

Still, the expected output of equilibrium under discretion is equal to its potential level, $\mathbb{E}[y_N^D] = \bar{y}$, and its variance is $\mathbb{V}[y_N^D] = \sigma_\epsilon^2$.

The results of the naive central banker under discretion are quite similar to those yielded by the stabilizer one. First, observe that there is an inflationary bias as well, and its magnitude also depends on both how higher is the social optimal level output compared to the long run level and the degree of conservativeness. The difference is in the absence of the output elasticity with respect to inflation surprise. Still, an ultra conservative central banker ($\phi = 0$) delivers the expected inflation equal to the inflation target. Second, the same can be said about the determinants of the inflation variance. Finally, given the assumption of rational expectations, inflation is completely anticipated by the agents, and the expected output is equal to its level of long run.

⁶To be more precise, the behavior of $\mathbb{E}[y_N^C] - \mathbb{E}[y_S^C]$ depends on the values of all parameters in the model. In fact, $\mathbb{E}[y_N^C] - \mathbb{E}[y_S^C] = a[a^2\bar{y}\lambda\phi - a\lambda(y^* - \bar{y}) + y^*\phi]$, such that its roots are $a = 0$ and $a = \left[\lambda(y^* - \bar{y}) \pm \sqrt{\lambda^2(y^* - \bar{y})^2 - 4\bar{y}y^*\lambda\phi}\right]/2\bar{y}\lambda\phi$, which can be either real (either a single or two different) or complex. In Figure D-1c, for example, we have the case of two different real roots ($\phi/\lambda = 1$ and $\phi/\lambda < 1$) and the case of complex roots ($\phi/\lambda > 1$).



We can compute the welfare of the naive central banker under discretion, which is given by

$$\begin{aligned} \mathbb{E} [L_N(\pi_N^D, \bar{y} + \epsilon)] &= \mathbb{E} [\max \{|\phi(y^* - \bar{y}) - \phi\epsilon|, \phi|\bar{y} - y^* + \epsilon|\}] \\ &= \phi \mathbb{E} [|y^* - \bar{y} - \epsilon|]. \end{aligned}$$

Clearly, we have $\mathbb{E} [L_N(\pi_N^D, \bar{y} + \epsilon)] > \mathbb{E} [L_N(\pi_N^C, y_N^C)]$, which implies that there is a worsening in the policymaker's welfare with relation to the setting under on-binding commitment.

Proposition 3. *Under discretion, the following statements are true:*

- (i) $\mathbb{E} [\pi_N^D] \geq \mathbb{E} [\pi_S^D]$ if and only if $\frac{\phi}{\lambda} \geq a$;
- (ii) $\mathbb{V} [\pi_N^D] \geq \mathbb{V} [\pi_S^D]$ if and only if $\frac{\phi}{\lambda} \geq a$.

It is possible to see Proposition 3 is based on the same condition of Proposition 1 and presents similar results. Therefore, the interpretation of itens (i) and (ii) of both are almost the same. The only difference is the fact of inflation and output now are yielded when policymakers act with discretion rather than with non-binding commitment. However, Proposition 3 no longer states convergence between expected inflations and inflation variances. In fact, as we can see in Figure D-1e in Appendix D, as a goes to infinity, $\mathbb{E} [\pi_N^C] - \mathbb{E} [\pi_S^C]$ goes to $-\infty$, that is, the inflation delivered by the stabilizer central banker becomes infinitely higher than one delivered by the naive one. The same is true about the variance. Mathematically, this difference is evident, because now we have monotonic functions while with commitment they present increasing and decreasing intervals.

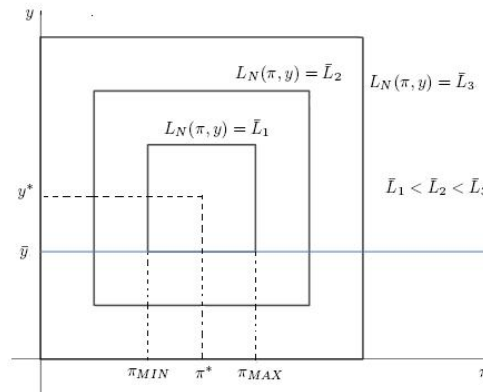
The reason for the absence of convergence between inflations under discretion is on the features of both central bankers' behavior. Recall that under discretion the expected inflation delivered by the stabilizer policymaker, given by (2), depends on the output's elasticity with respect to inflation surprise. As he chooses inflation taking the inflation expectation as given, the higher the elasticity, the more the marginal gain of the inflation surprise in terms of output. Thus, as a increases, his inflation is higher with the aim of creating inflation surprise. On the other hand, the inflation delivered by the naive policymaker does not depend on a , what we can see in (12). As this central banker does not improve his welfare by substituting price level by output, a higher elasticity would improve y but also would not changed π , what would let his welfare unchanged. Therefore, his incentive to create inflation surprise is not affected by changes in the elasticity.

4.2. Commitment

We have already seen the results of both central bankers under commitment, but it is worthy to take a deeper look in the naive behavior in this setting, given its singular features. Recall that, under commitment, the central banker's promise of pursuing the inflation target is credible, such that $\pi^e = \pi = \pi^*$ and there is no inflationary bias. In this case, the expected output is equal to its potential level as well. By keeping its promise, the monetary authority's expected loss is then $\mathbb{E} [L_N(\pi^*, \bar{y} + \epsilon)] = \phi \mathbb{E} [|y^* - \bar{y} + \epsilon|]$. In addition, the characteristics of the naive behavior make this same welfare be achieved for infinitely many other pairs (π, y) . In order to see this, observe that now the supply curve is just $y = \bar{y} + \epsilon$, such that it is represented by a horizontal line in the plane $\pi - y$.

As Figura 2 shows, the fact of the naive central banker has rectangles of indifference implies there is a line segment in which the supply curve coincides with the loss function's level curve. Therefore, there is a continuum of inflation levels which yields the same loss that the inflation target does. In fact, notice that inflation target is the center of the interval $[\pi_{\text{MIN}}, \pi_{\text{MAX}}]$. We must characterize it, especially with respect to its width and its determinants. In order to do so, we first have to find the value of its

Figure 2. Naive behavior under commitment.



endpoints. Note that $\pi_{\text{MIN}} < \pi$ and $\pi_{\text{MAX}} > \pi$, such that the first one fits item (i) and the second fits item (ii) of Proposition 2. Therefore, they must satisfy

$$\begin{aligned} \pi_{\text{MIN}} &= \pi^* - \phi(y^* - \bar{y} + \epsilon) \\ \pi_{\text{MAX}} &= \pi^* + \phi(y^* - \bar{y} + \epsilon) = \pi^D, \end{aligned}$$

where we use the fact of $y = \bar{y} < y^*$.

This first analysis, however, may be misleading, because it might indicate that the naive central banker can choose any inflation in the range $[\pi_{\text{MIN}}, \pi_{\text{MAX}}]$ without changing the slope of the supply curve and thus letting his welfare be unchanged. Recall that his optimization problem is constrained by (1), such that its slope changes whenever $\pi^e \neq \pi$. Therefore, if the central banker chooses $\pi \neq \pi^*$ and public believes that the commitment is binding, the supply curve satisfies (11), which is no longer a horizontal line. This implies that every deviation of his announcement has gain (or cost) in terms of welfare. In fact, any policy deviation makes the analysis of section 3 the most suitable, as we have seen. Nevertheless, the shape of the indifference curves of the naive central banker seems to indicate that small deviations might cause little changes in $L_N(\pi, y)$, given that the slope of (11) would change little as well. The next section studies the gain of the deviation and compares the results of the naive central banker with our benchmark case.

5. THE GAIN OF DEVIATION UNDER COMMITMENT

As we have seen, the shape of indifference curves of the naive central banker seems to indicate that small deviations from the commitment $\pi = \pi^*$ causes small changes in his welfare. However, it is necessary to use our benchmark case in order to define what we are calling “small changes”. In this section we use two measures to evaluate the incentives that both policymakers face to cheat the public. The first one is the optimal deviation: how much the central banker would deviate from the inflation target, once he has assumed the commitment of pursuing it, if he made his choice by maximizing his welfare.

Let $d_i = \pi - \pi^*$ be the deviation from the inflation target under commitment for $i = S, N$. Observe that we can write L_i in terms of d , which gives us the following expression for the stabilizer:

$$L_S(d) = \frac{1}{2} [d^2 + \lambda(\bar{y} - y^* + ad + \epsilon)^2].$$



This quadratic function reaches its minimum when $d_S = a\lambda[(y^* - \bar{y}) - \epsilon]/(1 + \lambda a^2)$, the same value of $\pi_S^C - \pi^*$. In other words, the optimal deviation for the stabilizer central banker is just the value of the inflationary bias yielded when he adopts π_S^C . Similarly, for the naive one we have $L_N(d) = \max\{|d|, \phi|\bar{y} - y^* + ad + \epsilon|\}$, which reaches its minimum when $d_N = \phi[(y^* - \bar{y}) - \epsilon]/(1 + \phi a)$. Once again, the optimal deviation is given by the bias obtained when the policymaker adopts his optimal policy under (non-binding) commitment, π_N^C .

There is no novelty in the above analysis: the naive central banker will choose an inflation which yields a higher deviation (bias) than one adopted by the stabilizer if and only if $\phi - \lambda a > 0$. We have already obtained such a result in section 3, which may make us to think that, whenever the mentioned condition is satisfied, the naive central banker has more incentive to deviate than the stabilizer. However, this measure accounts for the absolute magnitude of the central bankers' welfare. As the functional forms of the loss functions are different from each other, it is not possible to compare the values of L_S with L_N . We need a measure which considers welfare but not in absolute terms.

We can use the idea of elasticity in order to overcome the aforementioned difficult. We must answer the following question: under (non-binding) commitment, what is the percentage change in the central banker's welfare if inflation increases one percent? When we evaluate this elasticity at the point $\pi = \pi^*$, we have a measure of the incentive to deviate from the inflation target, independent of the absolute magnitude of the central banker's loss. Then, define

$$\eta_i = \left. \frac{\partial L_i}{\partial \pi} \frac{\pi_i^C}{L_i} \right|_{\pi_i^C = \pi^*}$$

as the elasticity, for $i = S, N$. Recall that the derivative is evaluated in $\pi^e = \pi^*$ and π_i^C is given by either (4) or (8).

Our results indicate that the stabilizer central banker has a welfare improvement higher than the naive one for the same 1% deviation from the inflation target:

$$\eta_S = -\frac{2a\pi^*}{y^* - \bar{y} - \epsilon}$$

$$\eta_N = -\frac{a\pi^*}{y^* - \bar{y} - \epsilon}.$$

Observe that the magnitude of the η_i is primarily determined by a and π^* : the more elastic the output to inflation surprise, the higher η_i is; similarly, the higher the inflation target, the higher η_i is. Furthermore, one can see that those elasticities do not depend on the conservativeness of the policymakers.

The conclusion of this section confirms what the graphical analysis of section 4 suggests, namely that the incentives for the naive central banker deviates from the commitment are lower than those of the stabilizer one. In fact, the result obtained here is even stronger: it states that, for the same 1% of deviation, the response (in percentage) of the welfare of the former is twice higher than the response of the latter. Moreover, this finding is independent of how conservative both policymakers are. Once again, the main reason for such a difference is on the nature of the naive behavior, since that it does not substitute price by output stabilization, which implies in low gains of creating inflation surprise.

6. CONCLUDING REMARKS

There is a case for a naive central banker. Whenever the supply curve's slope is large enough, the naive policymaker delivers both expected inflation and inflation variance lower than the stabilizer one. If both types present the same level of conservativeness ($\phi/\lambda = 1$), for example, such a condition is just that the output be inelastic with respect to inflation surprise. In order to have more economic intuition, we must then analyse what are the determinants of this elasticity. Given that we adopt the Lucas supply curve,

the underlying framework of our model is based on the assumption of information rigidity. Therefore, the more rigid the information, the slower the adjust process in the economy, and the more inelastic the output with respect to inflation surprise (the higher the supply curve's slope). Therefore, the case for the naive central banker requires that the economy presents enough information rigidity. In addition, we conjecture that this result does not change when one adopts assumption of price or wage stickiness.

Because rigidity is a relatively common feature of many modern economies, a natural conclusion would be that many central banks should be headed by a naive individual. Yet, there is no empirical evidence of this fact. In fact, it is only possible to find few researchers and politicians which recommend such an option. Our conjecture for this absence is that a naive central banker will in general be more populist than the stabilizer one. The reason is on the nature of the naive behavior, described in section 3, which is more subject to political pressures. Then, the more populist (the less conservative) the naive relative to the stabilizer, the more rigid the economy must be for the naive central banker delivers inflation lower than the stabilizer one. For example, as we have seen, if $\phi/\lambda = 1$, then the condition is $a < 1$, but if $\phi/\lambda = 4$, then the output should be four times more inelastic than the previous case. In short, the more populist the naive central banker, the "weaker" the case for him.

As the remainder literature on inflationary bias, our model can be extended by considering that the central banker chooses inflation indirectly through his instruments (money supply, for instance). This extension would improve the discussion about rules and discretion in monetary policy, virtually absent in our analysis. A richer modification in our original framework would allow us to build our results on a New Keynesian approach. By adopting such an alternative, we would be able to study the behavior of interest rate, inflation and output simultaneously. In addition, if the chosen model were a DSGE, one would be possible to consider long run effects as well as impacts of other shocks on the naive behavior. However, we believe that the main contribution of adopting such models is to overcome the limitation imposed by the lack of dynamics. By using the New Keynesian framework, it is possible to study reputation issues and how penalizations can affect the behavior of the naive central banker over time. We conjecture that this modification would make the results of both types of central bankers (naive and stabilizer) even more similar, because the effects of reputation incentives. Yet, this extension may not be so direct, given the non-smoothness of the loss function, but it would add realism on the monetary policy analysis and would make the results closer to the state of the art in monetary theory. Finally, our results would be enriched if we added issues concerning to labor market, for example, like those studies of Guzzo & Velasco (1999); Jerger (2002); Lippi (2002).

REFERENCES

- Backus, D., & Driffill, J. (1985). Inflation and reputation. *American Economic Review*, 75(3), 530–538.
- Ball, L. (1995). Time consistent inflation policy and persistent changes in inflation. *Journal of Monetary Economics*, 36(2), 329–350.
- Barro, R., & Gordon, D. (1983a). A positive theory of monetary policy in a natural rate model. *Journal of Political Economy*, 91, 589–610.
- Barro, R., & Gordon, D. (1983b). Rules, discretion, and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12, 101–121.
- Blinder, A. S. (2008). *The quiet revolution: Central banking goes modern*. Yale University Press.
- Blinder, A. S. (2010). How central should the central bank be? *Journal of Economic Literature*, 48(1), 123–133.
- Canzoneri, M. (1985). Monetary policy games and the role of private information. *American Economic Review*, 75(4), 1056–1070.
- Capie, F., Fischer, S., Goodhart, C., & Schnadt, N. (1995). *The future of central banking*. Cambridge Books.
- Chari, V., & Kehoe, P. (1990). Sustainable plans. *Journal of Political Economy*, 98(4), 783–802.
- Clarida, R., Galí, J., & Gertler, M. (1999). The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 37, 1661–1707.



- Cukierman, A., & Liviatan, N. (1991). Optimal accommodation by strong policymakers under incomplete information. *Journal of Monetary Economics*, 27(1), 99–127.
- Damjanovic, T., Damjanovic, V., & Nolan, C. (2008). Unconditionally optimal monetary policy. *Journal of Monetary Economics*, 55(3), 491–500.
- Epstein, G. (2005). Central banks as agents of economic development. *PERI Working Papers*, 103.
- Goodhart, C. A. E. (2011). The changing role of central banks. *Financial History Review*, 18(02), 135–154.
- Griebeler, M. d. C. (2015). On the existence of loss function for some useful classes of central bankers. *International Journal of Monetary Economics and Finance*, forthcoming.
- Guzzo, V., & Velasco, A. (1999). The case for a populist central banker. *European Economic Review*, 43(7), 1317–1344.
- Herrendorf, B., & Lockwood, B. (1997). Rogoff's "conservative" central banker restored. *Journal of Money, Credit, and Banking*, 476–495.
- Jerger, J. (2002). How strong is the case for a populist central banker? A note. *European Economic Review*, 46(3), 623–632.
- Jonsson, G. (1997). Monetary politics and unemployment persistence. *Journal of Monetary Economics*, 39(2), 303–325.
- Kydland, F. E., & Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *The Journal of Political Economy*, 473–491.
- Lippi, F. (2002). Revisiting the case for a populist central banker. *European Economic Review*, 46(3), 601–612.
- Nobay, A., & Peel, D. (2003). Optimal discretionary monetary policy in a model of asymmetric central bank preferences. *The Economic Journal*, 113, 657–665.
- Orphanides, A. (2012). New paradigms in central banking? *Chapters in SUERF Studies*, 13–32.
- Rogoff, K. (1985). The optimal commitment to an intermediate monetary target. *Quarterly Journal of Economics*, 100(4), 1169–1189.
- Ruge-Murcia, F. (2003). Inflation targeting under asymmetric preferences. *Journal of Money, Credit and Banking*, 35(5), 763–785.
- Sauer, S. (2010). When discretion is better: Initial conditions and the timeless perspective. *Economics Letters*, 107(2), 128–130.
- Voorneveld, M. (2014). From preferences to Leontief utility. *Economic Theory Bulletin*, 1–8.
- Walsh, C. (2010). *Monetary theory and policy* (3rd ed.). Cambridge: MIT Press.

APPENDIX A. THE STABILIZER CENTRAL BANKER'S OPTIMAL MONETARY POLICY

Discretion

The loss function of the stabilizer central banker as a function only of inflation is given by

$$L_S(\pi) = \frac{1}{2} \left[(\pi - \pi^*)^2 + \lambda(\bar{y} + a(\pi - \pi^e) + \epsilon - y^*)^2 \right],$$

which yields the following first order condition (FOC):

$$(\pi - \pi^*) + \lambda a(\bar{y} + a(\pi - \pi^e) + \epsilon - y^*) = 0. \tag{A-1}$$

By rearranging the terms of (A-1) and using the fact of $\pi^e = \pi$ we have

$$\begin{aligned} \pi_S^D &= \pi^* + \lambda a(y^* - \bar{y} - \epsilon) & \tag{A-2} \\ \mathbb{E}[\pi_S^D] &= \pi^* + \lambda a(y^* - \bar{y}) \\ \mathbb{V}[\pi_S^D] &= \lambda^2 a^* \sigma_\epsilon^2. \end{aligned}$$

By substituting (A-2) into (1) we have the economy's output:

$$\begin{aligned} y_S^D &= \bar{y} + (1 - \lambda a^2)\epsilon \\ \mathbb{E}[y_S^D] &= \bar{y} \\ \mathbb{V}[y_S^D] &= (1 - \lambda a^2)^2 \sigma_\epsilon^2. \end{aligned}$$

Commitment

We have already seen that when the commitment is binding $\pi = \mathbb{E}[\pi] = \pi^e = \pi^*$, which yields $y = \bar{y} + \epsilon$, $\mathbb{E}[y] = \bar{y}$ and $\mathbb{V}[y] = \sigma_\epsilon^2$. On the other hand, when the commitment is not binding, the central banker's optimization problem yields (A-1) as its FOC, but now $\pi^e = \pi^*$. Then, we have

$$\begin{aligned} \pi_S^C &= \pi^* + \frac{\lambda a}{1 + \lambda a^2} (y^* - \bar{y} - \epsilon) \tag{A-3} \\ \mathbb{E}[\pi_S^C] &= \pi^* + \frac{\lambda a}{1 + \lambda a^2} (y^* - \bar{y}) \\ \mathbb{V}[\pi_S^C] &= \frac{\lambda^2 a^2 \sigma_\epsilon^2}{(1 + \lambda a^2)^2}. \end{aligned}$$

By substituting (A-3) into the supply curve (1),

$$\begin{aligned} y_S^C &= \frac{\bar{y} + \lambda a^2 y^* + \epsilon}{1 + \lambda a^2} \\ \mathbb{E}[y_S^C] &= \frac{\bar{y} + \lambda a^2 y^*}{1 + \lambda a^2} \\ \mathbb{V}[y_S^C] &= \frac{\sigma_\epsilon^2}{(1 + \lambda a^2)^2}. \end{aligned}$$

The gain of the deviation under commitment

Let us calculate the elasticity η_S :

$$\begin{aligned} \eta_S &= \frac{(\pi_S^C - \pi^*) + \lambda a(\bar{y} + a(\pi_S^C - \pi^*) + \epsilon - y^*)\pi_S^C}{\frac{1}{2}[(\pi_S^C - \pi^*)^2 + \lambda(\bar{y} + a(\pi_S^C - \pi^*) + \epsilon - y^*)^2]} \Bigg|_{\pi_S^C = \pi^*} \\ &= \frac{\lambda a(\bar{y} + \epsilon - y^*)\pi^*}{\frac{1}{2}\lambda(\bar{y} + \epsilon - y^*)^2} \\ &= \frac{-2a\pi^*}{y^* - \bar{y} - \epsilon}. \end{aligned}$$

APPENDIX B. THE NAIVE CENTRAL BANKER'S OPTIMAL MONETARY POLICY

The steps of the results of commitment and discretion are shown in the text.



The gain of the deviation under commitment

By using Proposition 2, one can see that π^* is lower than π_N^C , such that it is in decreasing part of L_N (see (10)). Thus, $\partial L_N / \partial \pi = -a\phi$ and $L_N(\pi) = \phi \left| \bar{y} + a(\pi_N^C - \pi^*) + \epsilon - y^* \right|$, which yields

$$\begin{aligned} \eta_N &= \frac{-a\phi\pi_N^C}{\phi \left| \bar{y} + a(\pi_N^C - \pi^*) + \epsilon - y^* \right|} \Big|_{\pi_N^C = \pi^*} \\ &= \frac{-a\phi\pi^*}{\phi(y^* - \bar{y} - \epsilon)} \\ &= \frac{-a\pi^*}{(y^* - \bar{y} - \epsilon)}, \end{aligned}$$

where we use the fact of $y^* > \bar{y}$ and assume that ϵ is small, since it is a shock.

APPENDIX C. OMITTED PROOFS

Proof. **Proposition 1**

(i) $\mathbb{E}[\pi_N^C] \geq \mathbb{E}[\pi_S^C]$

$$\begin{aligned} \pi^* + \frac{\phi(y^* - \bar{y})}{1 + a\phi} &\geq \pi^* + \frac{\lambda a(y^* - \bar{y})}{1 + \lambda a^2} \\ \frac{\phi(y^* - \bar{y})}{1 + a\phi} &\geq \frac{\lambda a(y^* - \bar{y})}{1 + \lambda a^2} \end{aligned} \tag{C-4}$$

$$\begin{aligned} \phi + \phi\lambda a^2 &\geq \lambda a + \phi\lambda a^2 \\ \phi - \lambda a &\geq 0 \Leftrightarrow a \leq \frac{\phi}{\lambda}. \end{aligned} \tag{C-5}$$

(ii) $\mathbb{V}[\pi_N^C] \geq \mathbb{V}[\pi_S^C]$

$$\begin{aligned} \frac{\phi^2 \sigma_\epsilon^2}{(1 + a\phi)^2} &\geq \frac{\lambda^2 a^2 \sigma_\epsilon^2}{(1 + \lambda a^2)^2} \\ \frac{\phi^2}{(1 + a\phi)^2} &\geq \frac{\lambda^2 a^2}{(1 + \lambda a^2)^2}, \end{aligned}$$

which implies (C-4) and then (C-5).

(iii) $\mathbb{E}[y_N^C] \geq \mathbb{E}[y_S^C]$

$$\bar{y} + \frac{a\phi y^*}{1 + a\phi} \geq \frac{\bar{y} + \lambda a^2 y^*}{1 + \lambda a^2}$$

$$\frac{a}{(1 + a\phi)(1 + \lambda a^2)} [a\bar{y}(\lambda + a\phi) + y^*(\phi - \lambda a)] \geq 0,$$

such that a sufficient condition is $\phi - \lambda a \geq 0$, that is, $a \leq \phi/\lambda$.

(iv) By applying the L'Hôpital rule in (4), (5), (6), (8), (9), and (10) we can obtain

$$\begin{aligned} \lim_{a \rightarrow +\infty} \mathbb{E}[\pi_N^C] &= \lim_{a \rightarrow +\infty} \mathbb{E}[\pi_S^C] = \pi^* \\ \lim_{a \rightarrow +\infty} \mathbb{V}[\pi_N^C] &= \lim_{a \rightarrow +\infty} \mathbb{V}[\pi_S^C] = 0 \\ \lim_{a \rightarrow +\infty} \mathbb{E}[y_N^C] &= \bar{y} + y^* > y^* = \lim_{a \rightarrow +\infty} \mathbb{E}[y_S^C]. \end{aligned}$$

□

Proof. Proposition 2

Let π_1 and π_2 be the solutions of $\pi - \pi^* = \phi[\bar{y} - y^* + a(\pi - \pi^e) + \epsilon]$ and $\pi - \pi^* = -\phi[\bar{y} - y^* + a(\pi - \pi^e) + \epsilon]$, respectively. For given π^* , y^* , a , ϕ and π^e , it suffices to show that $|\pi_1 - \pi^*| > |\pi_2 - \pi^*|$, because $|\pi_i - \pi^*| = \phi|y_i - y^*|$ for $i = 1, 2$. We already computed π_2 when we solved the model with discretion, such that $\pi_2 = \pi^D$, given by (11). For π_1 , we have

$$\pi_1 - \pi^* = \phi[\bar{y} - y^* + a(\pi_1 - \pi^e) + \epsilon].$$

When one solves for π_1 , one obtains

$$\pi_1 = \frac{\pi^* - \phi[(y^* - \bar{y}) + a\pi^e + \epsilon]}{1 - a\phi}.$$

Then, by subtracting π^* of both sides and taking the absolute value, it results in

$$|\pi_1 - \pi^*| = \frac{|-\phi[(y^* - \bar{y}) + a(\pi^e - \pi^*) + \epsilon]|}{|1 - a\phi|}. \quad (\text{C-6})$$

For π_2 , observe that we can once again subtract π^* in both sides of (11) and then take the absolute value, such that

$$|\pi_2 - \pi^*| = \frac{|\phi[(y^* - \bar{y}) + a(\pi^e - \pi^*) + \epsilon]|}{|1 + a\phi|}. \quad (\text{C-7})$$

Now, given that $|x| = |-x|$ for all x , the numerators of (C-6) and (C-7) are equal. Still, recall that $\phi a > 0$, such that $|1 - \phi a| < |1 + \phi a|$. Therefore, $|\pi_2 - \pi^*| < |\pi_1 - \pi^*|$, which implies that $L(\pi_2, y_2) < L(\pi_1, y_1)$. □

Proof. Proposition 3

$$(i) \quad \mathbb{E}[\pi_N^D] \geq \mathbb{E}[\pi_S^D]$$

$$\pi^* + \phi(y^* - \bar{y}) \geq \pi^* + \lambda a(y^* - \bar{y})$$

$$\phi \geq \lambda a \Leftrightarrow \frac{\phi}{\lambda} \geq a.$$

$$(ii) \quad \mathbb{V}[\pi_N^D] \geq \mathbb{V}[\pi_S^D]$$

$$\phi^2 \sigma_\epsilon^2 \geq a^* \lambda^2 \sigma_\epsilon^2$$

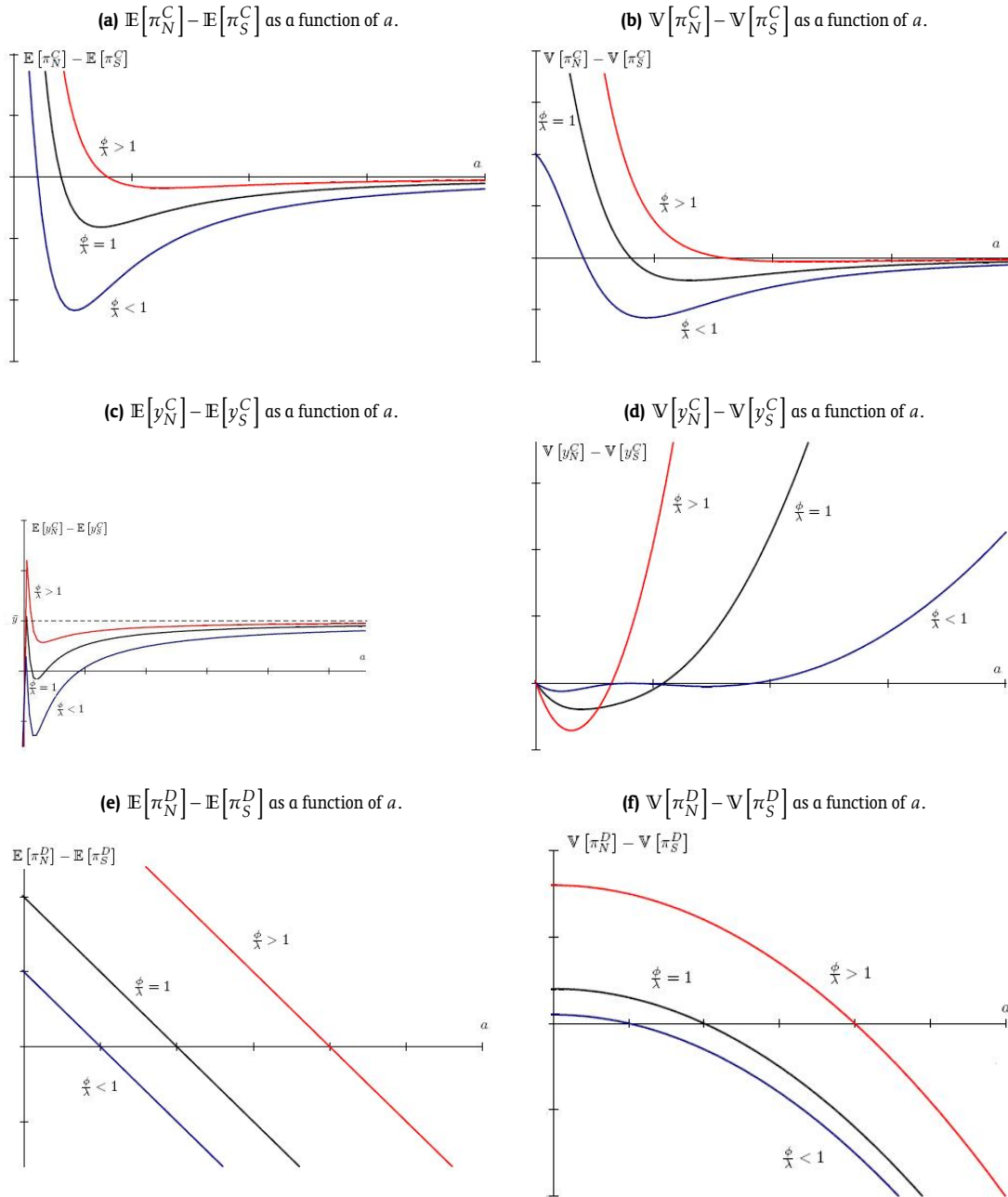
$$\frac{\phi}{\lambda} \geq a.$$

□



APPENDIX D. GRAPHICS

Figure D-1. Gráficos.



Testing the Optimality of Consumption Decisions of the Representative Household: Evidence from Brazil*

Marcos Gesteira Costa[†], Carlos Enrique Carrasco-Gutierrez[‡]

Contents: 1. Introduction; 2. Method of Estimation; 3. Data; 4. Empirical Results; 5. Conclusion.

Keywords: CCAPM, Rule of Thumb, Aggregate Consumption, Permanent Income Hypothesis, Euler Equations.

JEL Code: C32, E21.

This paper investigates whether there is a fraction of consumers that do not behave as fully forward-looking optimal consumers in the Brazilian economy. The generalized method of moments technique was applied to nonlinear Euler equations of the consumption-based capital assets model contemplating utility functions with time separability and non-separability. The results show that when the household utility function was modeled as constant relative risk aversion, external habits and Kreps–Porteus, estimates of the fraction of rule-of-thumb households was, respectively, 89%, 78% and 22%. According to this, a portion of disposable income goes to households who consume their current incomes in violation of the permanent income hypothesis.

Este artigo investiga se existe um comportamento subótimo nas decisões de consumo intertemporal na economia brasileira. O Método Generalizado de Momentos (MGM) foi aplicado ao modelo de precificação de ativos baseados no consumo (CCAPM) contemplando funções de utilidade separáveis e não separáveis. Os resultados mostram a existência de uma parcela da população que segue esse comportamento conhecido também como rule of thumb. Em termos mais concretos a fração de indivíduos que segue a regra de bolso é, respectivamente, de 89%, 78% e 22% para as funções de utilidade modeladas segundo as hipóteses de aversão relativa ao risco constante, hábitos externos e Kreps–Porteus. De acordo com esses resultados, uma grande parcela da população consome totalmente a sua renda corrente o que implica em violação da hipótese da renda permanente.

*This article is a revised version of Marcos Gesteira's Master Thesis, done under the supervision of Carlos Enrique Carrasco-Gutierrez. We gratefully acknowledge the comments and suggestions Fabio Reis Gomes, Wilfredo Maldonado, José Angelo Divino and seminar participants at UCB Brasília and Sociedade Brasileira de Econometria–SBE conference in Natal, 2014. We thank an anonymous referee and Alexandre B. Cunha (Editor) for the suggestions on an earlier version of this article.

[†]Universidade Católica de Brasília, SGAN 916, Módulo B, Brasília, DF, Brasil. E-mail: marcogest@hotmail.com

[‡]Universidade Católica de Brasília, SGAN 916, Módulo B, Brasília, DF, Brasil. CEP 70190-045. E-mail: carlosenrique@ucb.br



1. INTRODUCTION

The permanent income hypothesis (PIH), described by Friedman (1957), states that transitory changes in income have little effect on consumer spending, while permanent income is responsible for most of the variation in consumption. In his seminal work, Hall (1978) founded a new approach to study aggregate consumption. By using Euler equations for the optimal choice of a representative consumer, he showed that consumption should follow a random walk and argued that this holds in empirical applications, for instance that postwar U.S. data are consistent with this implication. In contrast, Flavin (1981), using a rational expectations structure, argued that consumption is sensitive to current income and it is greater than that predicted by the permanent income hypothesis. This conclusion has been widely interpreted as evidence of the existence of liquidity constraint. Empirical evidence shows that liquidity constraint is one of the main reasons why it is difficult to observe consumption smoothing in the data. Based on this evidence, Campbell & Mankiw (1989, 1990) suggested that aggregated data on consumption would be better characterized if there were two types of consumers. They nested the PIH in a more general model in which a proportion λ of consumers follow the rule of thumb,¹ consuming their current income (myopic spenders), while the remaining $(1 - \lambda)$ individuals consume optimally (optimizing savers). Using log-linearization of the model and instrumental variable estimates, they established by empirical application that there was a strong violation of the permanent income hypothesis because a significant fraction of the households have suboptimal behavior.

Cushing (1992) and Weber (2002) used intertemporally non-separable utility functions to study the behavior of American consumers. Cushing used a quadratic utility function modeled with current consumption and once-lagged consumption. Weber (2002) generalized Cushing's analysis by modeling the rule of thumb in nonlinear Euler equations and using the generalized method of moments (GMM) estimation technique. In particular, he tested if the lifetime utility function is time non-separable and concluded that the effect of the rule of thumb was small and not statistically significant.

In this article, we follow the insight of Weber, who considered that consumption of the optimizing agent is aggregate consumption minus rule-of-thumb consumption. In addition, we use the consumption-based asset pricing model (CCAPM) of Breeden (1979) and Lucas (1978) as a base of modeling and testing. The CCAPM setup considers not only an interest rate as studied in Hall (1978), Flavin (1981), Campbell & Mankiw (1989) or Weber (2002), but several assets in the economy. For instance, Hansen & Singleton (1982, 1983) developed and tested the empirical implications of the PIH when asset returns are time-varying and stochastic. They used the S&P 500 index and Treasury Bill yield as a risk-free rates of return. Epstein & Zin (1991) used five individual stock return indexes which give value-weighted returns for broad groups of industrial stocks and Treasury Bill yields as a risk-free rates of return.

Regarding Brazilian data, some authors have tested the PIH by incorporating rule-of-thumb behavior, but no one has used this procedure of testing in the CCAPM setup². Among the papers that have studied the rule-of-thumb proportion of consumption for the Brazilian economy are the articles of Cavalcanti (1993), Reis, Issler, Blanco, & de Carvalho (1998), Issler & Piqueira (2000), Gomes (2004) and Gomes & Paz (2004).

Cavalcanti (1993), studying the intertemporal elasticity of substitution with data from 1,980 to 1989, contemplated a budget constrained consumer in one of the models. He found that 32% of the population followed the rule of thumb. Reis et al. (1998), as Campbell & Mankiw (1989), used a model in which a portion of the population was restricted to consume only current income in order to test the validity of the PIH. Their study ranged from 1947 to 1994. The econometric tests revealed that about 80% of the population was restricted to consume only their current income. Issler & Piqueira (2000) conducted a study on the temporal series of consumption in Brazil from 1947 to 1994, aiming to examine theoretical

¹These consumers are restricted to consuming their current income, with no optimizing behavior

²Issler & Piqueira (2000) tested the implication of different consumption models in the CCAPM setup but did not test rule-of-thumb behavior.

issues of the PIH. The main results pointed to the acceptance of cointegration between consumption and income, and they also found that about 74% of the individuals are restricted in terms of liquidity. Gomes (2004) used Beveridge and Nelson's decomposition to disclose a cyclical component in consumption when testing the PIH. When he adopted the habit formation specification he found similar results to those of Reis et al. (1998).

Gomes & Paz (2004) used panel data to test the applicability of the Keynesian theory, the PIH and the hybrid model to consumption decisions for Argentina, Brazil, Chile, Colombia, Peru, Paraguay and Uruguay. They used data from 1951 to 2000, and found values of λ that ranged from 47% for Peru to 79% for Argentina. The fraction of Brazilian income that belongs to consumers constrained to spend their current income was 61% in their study. Arreaza (2000) examined whether liquidity constraints or voracity effects could explain consumption and saving in Latin America, using panel data from the period 1973 to 1993. She rejected the PIH, since around 20% of the consumption in Latin American follows predicted current income.

This article makes some contributions to the literature on aggregate consumption. First, we use a new procedure to test the rule-of-thumb behavior for Brazilian consumers. We use the consumption-based asset pricing model (CCAPM), which allows more than one interest rate. Second, this paper generalizes the rule-of-thumb model to allow intertemporal non-separability in the representative household's preferences considering external habits (Abel, 1990). In addition, the Kreps-Porteus (Epstein & Zin, 1989, 1991) expected lifetime utility function, which separates the coefficient of relative risk aversion from the intertemporal elasticity of substitution, is employed. As complementary analysis, we study the traditional utility functional forms of constant relative risk aversion (CRRA). These different types of utility functions permit estimating the structural parameters: intertemporal discount factor, intertemporal elasticity of substitution, relative risk aversion coefficient and the habit formation parameter. The purpose of this work is not to criticize the methods employed in previous articles, but instead to show a new procedure to test rule-of-thumb behavior for the Brazilian economy.

The empirical results in this paper provide evidences of rule-of-thumb behavior in the Brazilian case. In other words, there is a proportion of the individuals consuming their current income, and another group of individuals that consume optimally in each period. Therefore, there was a strong violation of the permanent income hypothesis.

The remainder of this paper is organized as follows. In section 2, the model with rule-of-thumb behavior is briefly discussed in the Euler equations for three different specifications. The estimation and results are detailed in section 3. Finally, the conclusions are in section 4.

2. METHOD OF ESTIMATION

2.1. Testing rule of thumb in the CCAPM framework

The idea behind the consumption-based capital assets model (CCAPM), established by Lucas (1978) and Breeden (1979), is that agents accumulate assets to ensure their future consumption plan, so the asset return series are related with the consumption series. The maximization problem faced by the agents is:

$$\max_{\{C_{2,t+s}, \theta_{t+s+1}\}_{s=0}^{\infty}} U_t(\cdot) \quad \text{s.t.} \quad \begin{cases} C_{2,t} + \theta_{t+1}P_t = \theta_t P_t + \theta_t d_t + Y_t \\ C_{2,t}, \theta_{t+1} \geq 0 \\ \text{and } \theta_0 \text{ is exogenous} \end{cases}, \quad (1)$$

where U_t is the utility function in period t ; $C_{2,t}$ is the aggregated household's consumption that consumes according to optimizing behavior; θ_t is a vector of the N assets; P_t is the assets' pricing vector for each period; and d_t is the assets' dividends vector.³ In each period, the agent receives an

³ $\theta_t P_t + \theta_t d_t$ is the total wealth the investor in period t , also called A_t ; $\theta_{t+1} P_t$ is the total wealth the investor will take from period t to period $t + 1$, also denoted A_{t+1} .



exogenous income Y_t , which is a state variable in the consumer problem. Solving this problem for $U_t = E_t [\sum_{s=0}^{\infty} \beta^s u(C_{2,t+s})]$ yields the Euler equations:

$$P_{j,t} = E_t \left[\beta \frac{\partial u_{t+1} / \partial C_{2,t+1}}{\partial u_t / \partial C_{2,t}} (P_{j,t+1} + d_{j,t+1}) \right], \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t, \quad (2)$$

where $u_t(\cdot)$ is the instantaneous utility function; β is the intertemporal discount coefficient; the index j refer to each available asset, and

$$\beta \frac{\partial u_{t+1} / \partial C_{2,t+1}}{\partial u_t / \partial C_{2,t}}$$

is the stochastic discount factor at $t + 1$. Dividing both sides by $P_{j,t}$ and placing the rights side under $(P_{j,t+1} + d_{j,t+1})$, it is possible to replace $(P_{j,t+1} + d_{j,t+1})/P_{j,t}$ by $R_{j,t+1}$, the gross return of asset j at $t + 1$, so that

$$1 = E_t \left[\beta \frac{\partial u_{t+1} / \partial C_{2,t+1}}{\partial u_t / \partial C_{2,t}} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (3)$$

Hall (1978), using a quadratic utility functional form and fixed return rate, reached the conclusion that the aggregate consumption series behaves as a random walk:

$$\Delta C_{2,t} = \epsilon_t, \quad (4)$$

where $\Delta C_{2,t}$ is the variation in consumption and ϵ_t was called innovation. Campbell & Mankiw (1989) divided consumers into two groups. The first group receives a share, λ , of the disposable income and consumes all their current income $Y_{1,t}$; the second group receives a share $(1 - \lambda)$ of the disposable income, follows the PIH and their income is $Y_{2,t}$. Hence, the total income of the economy is $Y_t = Y_{1,t} + Y_{2,t}$, or

$$Y_t = \lambda Y_t + (1 - \lambda) Y_t. \quad (5)$$

The consumers from the first group have $\Delta C_{1,t} = \Delta Y_{1,t} = \lambda \Delta Y_t$, while the consumers from the second group follow equation (4). The total variation in consumption can be stated as $\Delta C_t = \Delta C_{1,t} + \Delta C_{2,t}$, and replacing this yields Campbell and Mankiw's test equation:

$$\Delta C_t = \lambda \Delta Y_t + (1 - \lambda) \epsilon_t. \quad (6)$$

This equation says that the variation in consumption is a weighted average between the variation of the income of the first group and the unpredictable variation in the permanent income of the second group. They specified their hypotheses as

$$\begin{aligned} H_0 : \Delta C_t = \epsilon_t, \text{ hence } \lambda = 0 \\ H_1 : \Delta C_t = \Delta Y_{1,t} = \lambda \Delta Y_t, \text{ hence } \lambda > 0. \end{aligned} \quad (7)$$

When $\lambda = 0$, the permanent income hypothesis holds. Under the alternative hypothesis the change in consumption is a weighted average of changes in current income. Equation (6) should not be estimated by ordinary least squares (OLS) since the error component may be correlated with changes in income.

Weber (2002) modeled consumption in nonlinear Euler equations, by isolating the consumption of the second group, $C_{2,t}$.⁴ So, let $C_t = C_{1,t} + C_{2,t}$, then $C_{2,t} = C_t - C_{1,t}$, and $C_{1,t} = \lambda Y_t$, then

$$C_{2,t} = C_t - \lambda Y_t. \quad (8)$$

⁴ $C_{2,t}$ is the consumption of the second group, optimizers. The consumers of the first group follow the rule of thumb, so their consumption cannot be modeled by the CCAPM. Note that we already use this notation in the maximization problem (equation (1)).

The Euler equations of the CCAPM problem are only valid for optimizing consumers, replacing (8) in equation (3), and yields

$$E_t \left[\beta \frac{u'(C_{t+1} - \lambda Y_{t+1})}{u'(C_t - \lambda Y_t)} R_{j,t+1} \right] = 1 \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (9)$$

Equation (9) can be used to estimate and to test the parameters of the model by the GMM technique. GMM estimators were developed by Peter Hansen in 1982. Since then this technique has enabled several breakthroughs in macroeconomics and finance research. The essence behind the GMM is to find a sample moment as close as possible to the population moment. Let $\tilde{\theta}$ be a vector of parameters and h an $r \times 1$ vector, where the lines are the orthogonality conditions. Then $\tilde{\theta}$ satisfies $E[h(\tilde{\theta}, w_t)] = 0$. In application with two assets' returns, the number of orthogonality conditions are $r = 2M$, where M is the number of instruments to be used in estimation. Let X_t be a vector of chosen instruments. Then the orthogonality conditions are

$$h(\tilde{\theta}, w_t) = \begin{bmatrix} \left(1 - \beta \frac{u'(C_{t+1} - \lambda Y_{t+1})}{u'(C_t - \lambda Y_t)} R_{1,t+1}\right) X_t \\ \left(1 - \beta \frac{u'(C_{t+1} - \lambda Y_{t+1})}{u'(C_t - \lambda Y_t)} R_{2,t+1}\right) X_t \end{bmatrix}_{2M \times 1}. \quad (10)$$

Therefore $E\left[\left(1 - \beta \frac{u'(C_{t+1} - \lambda Y_{t+1})}{u'(C_t - \lambda Y_t)} R_{j,t+1}\right) \otimes X_t\right] = 0$, for $j = 1, 2$. The sample moment is defined as $g(\tilde{\theta}, y_t) = \frac{1}{T} \sum_{t=1}^T h(\tilde{\theta}, w_t)$ and the GMM's estimator $\tilde{\theta}$ is the one that minimizes the scalar $Q(\tilde{\theta}, y_t) = [g(\tilde{\theta}, y_t)]' W [g(\tilde{\theta}, y_t)]$, where W is the weighting matrix which acts to weight the various moment conditions to build the distance measure.

A test for the over-identifying restrictions (TJ-test) allows checking whether the model's moment conditions match the data well or not. The TJ statistic employed is asymptotically chi-squared with $r - k$ degrees of freedom, where r is the number of orthogonality conditions and k the number of parameters in the structural model.

2.2. Utility functions

The utility's functional forms Constant Relative Risk Aversion Preferences (CRRA), external habits and Kreps–Porteus address time separability and non-separability.

The Constant Relative Risk Aversion Preferences

In the first model, the instantaneous utility function is parameterized as

$$u(C_{2,t}) = \frac{C_{2,t}^{1-\gamma} - 1}{1-\gamma},$$

and the utility function U_t is

$$U_t = E_t \left[\sum_{s=0}^{\infty} \beta^s u(C_{2,t+s}) \right] = E_t \left[\sum_{s=0}^{\infty} \beta^s \left\{ \frac{C_{2,t+s}^{1-\gamma} - 1}{1-\gamma} \right\} \right], \quad (11)$$

where γ is the relative risk aversion coefficient and the reciprocal of the consumption's intertemporal elasticity of substitution $\psi = 1/\gamma$.

The Euler equations are

$$1 = E_t \left[\beta \left(\frac{C_{2,t+1}}{C_{2,t}} \right)^{-\gamma} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (12)$$



Replacing (8) in (12), yields

$$1 = E_t \left[\beta \left(\frac{C_{t+1} - \lambda Y_{t+1}}{C_t - \lambda Y_t} \right)^{-\gamma} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (13)$$

Stationary regressors are obtained dividing through C_t , therefore

$$1 = E_t \left[\beta \left(\frac{\frac{C_{t+1}}{C_t} - \lambda \frac{Y_{t+1}}{C_t}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{-\gamma} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (14)$$

Let X_t be a vector of chosen instruments, thus the orthogonality conditions are

$$E \left[\left(1 - \beta \left(\frac{\frac{C_{t+1}}{C_t} - \lambda \frac{Y_{t+1}}{C_t}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{-\gamma} R_{j,t+1} \right) \otimes X_t \right] = 0 \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (15)$$

The External Habits Preferences

This parametric form of the individual preference assume that individual keeps the history of her own consumption, viewed as consumer's habit, allowing for non-separability of the utility function over time. The instantaneous utility function for External Habits used is

$$u(C_{2,t}, v_t) = \frac{\left[\frac{C_{2,t}}{v_t} \right]^{1-\gamma}}{1-\gamma}.$$

Following Abel (1990), we specify the function $v_t(\cdot)$ here as $v_t = \left[C_{2,t-1}^D \bar{C}_{2,t-1}^{1-D} \right]^\kappa$. In order to have "external habit", we set $D = 0$ and $\kappa > 0$. Therefore $v_t = \left[\bar{C}_{2,t-1} \right]^\kappa$ and the utility function U_t is

$$U_t = E_t \left[\sum_{s=0}^{\infty} \beta^s u(C_{2,t+s}, v_{t+s-1}) \right] = E_t \left[\sum_{s=0}^{\infty} \beta^s \left[\frac{\left[\frac{C_{2,t+s}}{\left(\bar{C}_{2,t+s-1} \right)^\kappa} \right]^{1-\gamma}}{1-\gamma} \right] \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t, \quad (16)$$

where $C_{2,t}$ is the individual consumption at t ; $\bar{C}_{2,t-1}$ is the per capita aggregated consumption at $t-1$; κ is a parameter controlling the time separability in the function.

The Euler equations are

$$1 = E_t \left[\beta \left(\frac{C_{2,t+1}}{C_{2,t}} \right)^{-\gamma} \left(\frac{C_{2,t}}{\bar{C}_{2,t-1}} \right)^{\kappa(\gamma-1)} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (17)$$

Replacing (8) in (17), yields

$$1 = E_t \left[\beta \left(\frac{C_{t+1} - \lambda Y_{t+1}}{C_t - \lambda Y_t} \right)^{-\gamma} \left(\frac{C_t - \lambda Y_t}{C_{t-1} - \lambda Y_{t-1}} \right)^{\kappa(\gamma-1)} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (18)$$

Stationary regressors are obtained dividing through C_t and C_{t-1} , therefore

$$1 = E_t \left[\beta \left(\frac{\frac{C_{t+1}}{C_t} - \lambda \frac{Y_{t+1}}{C_t}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{-\gamma} \left(\frac{\frac{C_t}{C_{t-1}} - \lambda \frac{Y_t}{C_{t-1}}}{1 - \lambda \frac{Y_{t-1}}{C_{t-1}}} \right)^{\kappa(\gamma-1)} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (19)$$

Representing in the unconditional form,

$$E \left[\left[1 - \beta \left(\frac{C_{t+1} - \lambda \frac{Y_{t+1}}{C_t}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{-\gamma} \left(\frac{C_t - \lambda \frac{Y_t}{C_{t-1}}}{1 - \lambda \frac{Y_{t-1}}{C_{t-1}}} \right)^{\kappa(\gamma-1)} R_{j,t+1} \right] \otimes X_t \right] = 0 \quad \text{for } j=1,2,\dots,N \quad \text{and } \forall t. \quad (20)$$

The Kreps–Porteus Preferences

The third utility preference treated here follow the Epstein & Zin (1989), being a generalization of the utility function proposed by Kreps & Porteus (1978). The aggregating function is parameterized as a constant elasticity of substitution (CES) function:

$$U_t = \left[(1 - \beta) C_{2,t}^\rho + \beta \left(E_t \tilde{U}_{t+1}^\alpha \right)^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}}, \quad \text{for } 0 \neq \rho < 1, \quad (21)$$

where E_t is the conditional expectation operator given the information available to the agent in the planning period and \tilde{U}_{t+1} is the agent's future utility. The consumption's intertemporal elasticity of substitution is $\psi = 1/(1 - \rho)$. The relative risk aversion coefficient γ is constant, $\gamma = 1 - \alpha$ where the parameter α reflects the agent's behavior towards risk. In particular, when $\alpha = 0$, we are back to the expected utility function with logarithmic preference. When $\alpha = \rho$, we have and additively separable utility function.

The Euler equations are

$$1 = E_t \left[\beta^\eta \left(\frac{C_{2,t+1}}{C_{2,t}} \right)^{\eta(\rho-1)} \tilde{B}_{t+1}^{\eta-1} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t, \quad (22)$$

where $\eta = \alpha/\rho$ and \tilde{B}_{t+1} is the optimum portfolio's gross return.

Replacing (8) in (22) yields

$$1 = E_t \left[\beta^\eta \left(\frac{C_{t+1} - \lambda Y_{t+1}}{C_t - \lambda Y_t} \right)^{\eta(\rho-1)} \tilde{B}_{t+1}^{\eta-1} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (23)$$

Stationary regressors are obtained dividing through C_t , therefore

$$1 = E_t \left[\beta^\eta \left(\frac{C_{t+1} - \lambda \frac{Y_{t+1}}{C_t}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{\eta(\rho-1)} \tilde{B}_{t+1}^{\eta-1} R_{j,t+1} \right] \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (24)$$

In the unconditional form representation we have:

$$E \left[\left[1 - \beta^\eta \left(\frac{C_{t+1} - \lambda \frac{Y_{t+1}}{C_t}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{\eta(\rho-1)} \tilde{B}_{t+1}^{\eta-1} R_{j,t+1} \right] \otimes X_t \right] = 0 \quad \text{for } j = 1, 2, \dots, N \quad \text{and } \forall t. \quad (25)$$

3. DATA

The series used for the estimation were the real per capita household consumption, real per capita gross domestic product (GDP), real gross returns of risky assets, and real gross returns of the riskless asset. The series of aggregate consumption, GDP and population are available at the website of IPEA (*Instituto de Pesquisa Economica Aplicada*), while the series of return on assets and rates of inflation are posted at the website of the Central Bank of Brazil.



The data range from 1995.Q1 to 2011.Q2. This period starts with the implementation of the *Plano Real*, the plan the Brazilian government launched that finally managed to end the persistently high inflation (with bouts of hyperinflation) that had held sway over the previous two decades. Another factor that contributed to this choice was that Reis et al. (1998) and Gomes (2004) suggested that the high value of λ they found was due to the credit constraint the Brazilian population encountered (high and unpredictable inflation with indexation not necessarily matched with salary indexation, making debt service as a proportion of household income extremely volatile). The lower inflation rates through the period studied in this paper resulted in credit expansion, the availability of funding to finance consumption was not at the same level as in the developed countries but was much higher than in the periods of the others studies. Therefore, a smaller part of the population following the rule of thumb was expected.

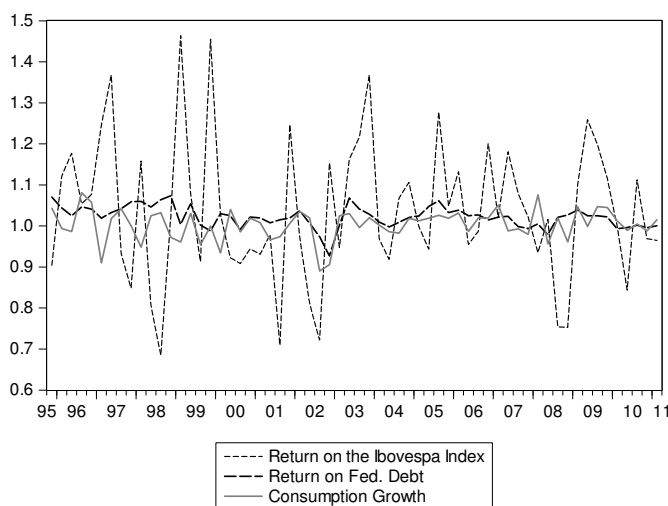
The series of the household consumption was calculated the same way as in Reis et al. (1998), where the gross fixed investment and current account balance series were subtracted from the GDP series to obtain a consumption of non-durable goods series.⁵

The returns of the IBOVESPA index were used to represent the returns of risky assets, because it is the most important index of average returns of the Brazilian stock market. Another interesting option would be IBrX, an index comprising more stocks that is widely used in the financial market for the static CAPM. However, the IBOVESPA series is longer and more suitable for the studied period. In order to represent the returns of the riskless asset in the Brazilian economy, the rate paid on government debt (SELIC rate) was used.⁶ The general price index (IGP-DI) calculated by *Fundação Getulio Vargas* (FGV) was used to deflate income, consumption and returns of both assets. The consumption and income data were also subject to seasonal adjustments. Figura 1 shows the data in quarterly frequencies.

4. EMPIRICAL RESULTS

In this section, we present GMM estimates of the rule-of-thumb models for the utility preferences shown in the previous section. In order to estimate the orthogonality conditions, generated by the Euler equa-

Figure 1. Data on Quarterly Frequency (1995.Q1 to 2011.Q2).



⁵Income and consumption series were divided by the population series. Linear interpolation and extrapolation were applied to transform an annual population series into a quarterly basis series.

⁶The regular savings account return rate was ruled out because its return rate is calculated by adding a fixed return to an inflation index, hence there would be no co-movement between the real returns of the regular savings account rate and consumption.

tions, we use several sets of instruments. The instruments correspond to lagged values of the growth in consumption and real interest rate.⁷ We also use the tests of the overidentifying restrictions (TJ-test) to assess the joint validity of each model and the set of instruments. In this paper, several sets of instruments were tested and none were rejected at the 5% level.

The results are presented in Tables 1, 2 and 3, where only display those where the parameter λ estimate was between 0 and 1.

For the CRRA utility, Tabela 1, the median estimate of the parameter λ was $\hat{\lambda} = 0.8945$ and all but one of them were significant at the 5% level.

That is, results for λ show that around 89% of the population follows the rule of thumb. All the estimates of the intertemporal discount coefficient, β , were significant at the 5% level, and their median was $\hat{\beta} = 0.9783$. The median for the relative risk aversion coefficient was $\hat{\gamma} = -0.0974$.

Tabela 2 reports the findings of the estimation of Euler equations (20) which correspond for the external habits utility model. Results show that overall estimations of the parameter λ for the most part were significant at the 5% level. The median of all valid estimates was $\hat{\lambda} = 0.7817$. It shows that around 78% of the Brazilian population follows the rule of thumb. All the results for the intertemporal discount coefficient, β , were significant at the 5% level and their median was $\hat{\beta} = 0.9793$. For the parameter κ this study found $\hat{\kappa} = 0.1518$, but only three of the estimates were significant at the 5% level. One positive feature of this κ estimation is that it does not violate the external habit basic assumption: if $\kappa > 0$ than $\gamma > 0$ or if $\kappa < 0$ than $\gamma < 0$. The relative risk aversion coefficient estimate was $\hat{\gamma} = 0.0548$ and only four of them were significant at the 5% level and only one at the 10% level.

For the Kreps–Porteus utility function model, the findings for the estimation of the Euler equations (25) are shown in Tabela 3. We found $\hat{\lambda} = 0.2263$, but in almost all cases λ estimates are not significant at the 5% level. The intertemporal discount coefficient, β , was significant at the 5% level in 21 estimations, and the median of all estimations was $\hat{\beta} = 0.9743$. Almost all the findings for $\eta = \alpha/\rho$ were very close to zero,⁸ however all the estimates, but one, were significantly different than zero. The median of all estimates was $\hat{\eta} = -0.00017$, which yields $\hat{\alpha} = 0.00013$ and $\hat{\gamma} = 1.00013$. For the parameter ρ , six estimates were significant at the 5% level and the median of all estimates was $\hat{\rho} = 0.7721$. The intertemporal elasticity of was captured by solving for each estimate $\psi = 1/(1 - \rho)$ and the median of all estimates was $\hat{\psi} = 3.5798$.⁹

In Tabela 4 we compare the estimates for β , γ and ψ with studies that also used the CCAPM framework but did not contemplate the rule of thumb parameter in the model, such as Issler & Piqueira (2000), Bonomo & Domingues (2002), Catalão & Yoshino (2006).¹⁰ Tabela 4 shows the results of these studies for β , γ and ψ .

The results for the intertemporal discount coefficient, β , were all significant at the 5% level, for the three functional forms and in line with the previous studies.

Comparing the relative risk aversion results to the previous studies, we obtained lower values for the CRRA and the external habits utility models. Both models' estimates were very close to zero. For the Kreps–Porteus utility function, the median of all estimates of the risk aversion parameter was $\hat{\gamma} = 1.00013$, in contrast what was found in Issler & Piqueira $\hat{\gamma} = 0.68$ and Bonomo & Domingues $\hat{\gamma} = 3.226$. The intertemporal elasticity of substitution was captured by solving for each estimate $\psi = 1/(1 - \rho)$ and the median of all estimates was $\hat{\psi} = 3.5798$, while estimates by Issler & Piqueira, and Bonomo & Domingues (2002) are $\hat{\psi} = 0.29$ and $\hat{\psi} = 0.371$, respectively.

⁷ A large number of instruments or a high number of assets can cause problems to find the optimal weighting matrix or influence the quality of asymptotic approximation, therefore the data must meet the following condition: $\frac{NM(NM+1)}{2} < NT$, where N are the number of Euler equations and T are the number of observations (Driscoll & Kraay, 1998). In this study, the number of instruments in the worst case the relation is $78 < 122$.

⁸ When $\alpha = 0$ and $\gamma = 0$, the preferences are logarithmic and it is not possible to test the separability in intertemporal consumption.

⁹ If the intertemporal elasticity of substitution was reached through $\hat{\psi} = 1/(1 - \hat{\rho})$, the estimate would be $\psi = 4.3879$.

¹⁰ The period of their estimations was from 1994 to 2003.



Table 1. Euler Equations for the CRRA Utility Function with Rule of Thumb.

$$E \left[\left(1 - \beta \left(\frac{C_{t+1} - \lambda \frac{Y_{t+1}}{C_t}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{-\gamma} R_{j,t+1} \right) \otimes X_t \right] = 0, \text{ for } j = 1, 2$$

where $N = 2$, $R_{1,t+1}$ = Ibovespa returns, and $R_{2,t+1}$ = Returns on Selic.

Inst./Mtx	β	γ	λ	P-value ($T \times J$)
I3/ASI	0.984*** (0.052)	0.944 (0.7365)	0.1530 (1.1568)	0.414
I5/ASI	0.981*** (0.0035)	0.050 (0.0517)	0.8049*** (0.0246)	0.462
I1/NWFSI	0.974*** (0.0041)	-0.974*** (0.0263)	0.8906*** (0.0006)	0.255
I4/NWFSI	0.979*** (0.0029)	-0.072 (0.0435)	0.8945*** (0.0099)	0.806
I3/NWFSI	0.978*** (0.0035)	-0.138 (0.1198)	0.8992*** (0.0204)	0.560
I6/NWFSI	0.971*** (0.0032)	-0.240 (0.3293)	0.9522*** (0.1109)	0.252
I4/NWVSI	0.969 (0.0004)	-1.275*** (0.0572)	0.9853*** (0.0042)	0.991
Median estimates	0.9783	-0.0974	0.8945	
Confidence interval for the median: $0.87 \leq \lambda \leq 0.915$				

Notes: (i) *, ** and *** denote, respectively, significance of parameter by the t -test at the 10%, 5% and 1% levels. (ii) The number in parentheses are the respective standard-deviation estimates, robust to heteroscedasticity and to serial correlation. (iii) The last line of the table shows the median of all estimates. (iv) List of instruments: I1 uses $R_{2,t}$, $R_{2,t-1}$, C_t/C_{t-1} , C_{t-1}/C_{t-2} ; I3 uses $R_{2,t-1}$, $R_{2,t-2}$, C_{t-1}/C_{t-2} , and C_{t-2}/C_{t-3} ; I4 uses $R_{2,t-1}$, $R_{2,t-2}$, $R_{1,t-1}$, $R_{1,t-2}$, C_{t-1}/C_{t-2} , and C_{t-2}/C_{t-3} ; I5 uses $R_{1,t-1}$, $R_{1,t-2}$, C_{t-1}/C_{t-2} , and C_{t-2}/C_{t-3} ; I6 uses $R_{2,t}$, $R_{2,t-1}$, $R_{1,t}$, $R_{1,t-1}$, C_t/C_{t-1} , and C_{t-1}/C_{t-2} . (v) The p -value of Hansen's overidentifying restrictions test results are shown in the last column. (vi) The table only reports the results that reached convergence in less than 1,000 iterations and was not rejected by Hansen's (1982) test. (vii) The first column shows the method by which the weighting matrix was reached, ASI denotes Andrews (1991), NWFSI denotes Newey & West (1987) with fixed windows, and NWVSI denotes Newey & West (1994), with variable windows.

Table 2. Euler Equations for the External Habits' Utility Function with Rule of Thumb.

$$E \left[\left(1 - \beta \left(\frac{C_{t+1} - \lambda \frac{Y_{t+1}}{C_t}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{-\gamma} \left(\frac{C_t - \lambda \frac{Y_t}{C_{t-1}}}{1 - \lambda \frac{Y_{t-1}}{C_{t-1}}} \right)^{\kappa(\gamma-1)} R_{j,t+1} \right) \otimes X_t \right] = 0, \text{ for } j = 1, 2$$

where $N = 2$, $R_{1,t+1} =$ Ibovespa returns, and $R_{2,t+1} =$ Returns on Gov.Debt-Selic.

Inst./Mtx	β	γ	λ	κ	P -value ($T \times J$)
I6/ASI	0.986*** (0.0031)	0.689** (0.259)	0.122 (0.5302)	0.582 (0.4940)	0.398
I5/NWFSI	0.626*** (0.2216)	10.655** (5.0581)	0.419 (0.2719)	1.6164*** (0.4853)	0.545
I5/ASI	0.467*** (0.2111)	10.111*** (7.9961)	0.594** (0.2240)	1.280*** (0.4685)	0.784
I4/ASI	0.977*** (0.0034)	0.067 (0.0921)	0.769*** (0.0861)	0.1540 (0.1658)	0.248
I4/NWFSI	0.979*** (0.0021)	0.054 (0.0632)	0.775*** (0.0644)	0.1517 (0.1245)	0.497
I8/NWFSI	0.978*** (0.0034)	0.035 (0.0734)	0.7784*** (0.0594)	0.1978 (0.1711)	0.149
I7/NWFSI	0.979*** (0.0033)	0.011 (0.0906)	0.7817*** (0.0813)	0.2008 (0.2471)	0.550
I6/NWVSI	0.977*** (0.0030)	0.120* (0.0669)	0.7956*** (0.0272)	0.0475 (0.0327)	0.374
I3/NWFSI	0.983*** (0.0031)	0.0579 (0.0523)	0.7983*** (0.0355)	0.0065 (0.0214)	0.542
I3/ASI	0.981*** (0.0034)	0.0458 (0.0526)	0.8063*** (0.0219)	0.0057 (0.0239)	0.330
I2/ASI	0.979 (0.0048)	-0.065** (0.0272)	0.8905*** (0.0008)	-0.0283** (0.0134)	0.284
I2/NWFSI	0.972 (0.0044)	-0.262 (0.3563)	0.9177*** (0.0817)	-0.1206 (0.1859)	0.287
Median estimates	0.9793	0.0548	0.7817	0.1518	
Confidence interval for the median: $0.717 \leq \lambda \leq 0.846$					

Notes: (i) *, ** and *** denote, respectively, significance of parameter by the t -test at the 10%, 5% and 1% levels. (ii) The number in parentheses are the respective standard-deviation estimates, robust to heteroscedasticity and to serial correlation. (iii) The last line of the table shows the median of all estimates. (iv) List of instruments: I2 uses $R_{1,t}$, $R_{1,t-1}$, C_t/C_{t-1} , C_{t-1}/C_{t-2} ; I3 uses $R_{2,t-1}$, $R_{2,t-2}$, C_{t-1}/C_{t-2} , and C_{t-2}/C_{t-3} ; I4 uses $R_{2,t-1}$, $R_{2,t-2}$, $R_{1,t-1}$, $R_{1,t-2}$, C_{t-1}/C_{t-2} , and C_{t-2}/C_{t-3} ; I5 uses $R_{1,t-1}$, $R_{1,t-2}$, C_{t-1}/C_{t-2} , and C_{t-2}/C_{t-3} ; I6 uses $R_{2,t}$, $R_{2,t-1}$, $R_{1,t}$, $R_{1,t-1}$, C_t/C_{t-1} , and C_{t-1}/C_{t-2} ; I7 uses $R_{2,t}$, $R_{2,t-1}$, $R_{2,t-2}$, $R_{1,t}$, $R_{1,t-1}$, and $R_{1,t-2}$; I8 uses $R_{2,t-2}$, $R_{2,t-3}$, $R_{1,t-2}$, and $R_{1,t-3}$. (v) The p -value of Hansen's overidentifying restrictions test results are shown in the last column. (vi) The table only reports the results that reached convergence in less than 1,000 iterations and was not rejected by Hansen's (1982) test. (vii) The first column shows the method by which the weighting matrix was reached, ASI denotes Andrews (1991), NWFSI denotes Newey & West (1987) with fixed windows, and NWVSI denotes Newey & West (1994), with variable windows.



Table 3. Euler Equations for the Kreps–Porteus Utility Function with Rule of Thumb.

$$E \left[\left(1 - \beta^\eta \left(\frac{C_{t+1} - \lambda \frac{Y_{t+1}}{C_{t+1}}}{1 - \lambda \frac{Y_t}{C_t}} \right)^{\eta(\rho-1)} \bar{B}_{t+1}^{\eta-1} R_{j,t+1} \right) \otimes X_t \right] = 0, \text{ for } j = 1, 2$$

where $N = 2$, \bar{B}_{t+1} = Ibovespa returns, $R_{2,t+1}$ = Returns on Gov.Debt-Selic, and

$$\eta = \frac{\alpha}{\rho}, \quad \psi = \frac{1}{1-\rho}, \quad \gamma = 1 - \alpha$$

Inst./Mtx	β	γ	λ	ρ	ψ	<i>P</i> -value ($T \times J$)
I3/ASI	0.9679*** (0.0048)	1.00006** (0.21349)	0.7692*** (0.1685)	0.8759*** (0.2775)	8.0590 (18.0265)	0.2984
I5/ASI	0.9733*** (0.0087)	1.000*** (0.03566)	0.0314 (1.4838)	-0.3779 (1.1368)	0.7257 (0.5987)	0.1146
I6/ASI	0.9743*** (0.0042)	1.0001*** (0.05109)	0.2263 (0.5098)	0.5219** (0.2257)	2.0919** (0.9880)	0.3767
I6/NWFSI	0.9772*** (0.0036)	1.0003*** (0.00507)	0.0297 (0.6660)	0.5438*** (0.1708)	2.1923*** (0.8212)	0.5140
I7/NWFSI	0.9751*** (0.00375)	1.0001*** (0.0632)	0.6572* (0.35836)	0.8127*** (0.2419)	5.3409 (6.9020)	0.3917
I1/NWVSI	0.9738*** (0.00356)	1.0000*** (0.0632)	0.1048 (4.5702)	0.9136*** (0.1729)	11.5762 (23.1740)	0.4359
I7/NWVSI	0.9768*** (0.00368)	1.0002*** (0.14673)	0.5843 (0.3813)	0.7220*** (0.2511)	3.5978 (3.2504)	0.5225
Median estimates	0.9793	0.0548	0.7817	0.1518		
Confidence interval for the median: $0.717 \leq \lambda \leq 0.846$						

Notes: (i) *, ** and *** denote, respectively, significance of parameter by the *t*-test at the 10%, 5% and 1% levels. (ii) The number in parentheses are the respective standard-deviation estimates, robust to heteroscedasticity and to serial correlation. (iii) The last line of the table shows the median of all estimates. (iv) List of instruments: I1 uses $R_{2,t}$, $R_{2,t-1}$, C_t/C_{t-1} , C_{t-1}/C_{t-2} ; I3 uses $R_{2,t-1}$, $R_{2,t-2}$, C_{t-1}/C_{t-2} , and C_{t-2}/C_{t-3} ; I5 uses $R_{1,t-1}$, $R_{1,t-2}$, C_{t-1}/C_{t-2} , and C_{t-2}/C_{t-3} ; I6 uses $R_{2,t}$, $R_{2,t-1}$, $R_{1,t}$, $R_{1,t-1}$, C_t/C_{t-1} , and C_{t-1}/C_{t-2} ; I7 uses $R_{2,t}$, $R_{2,t-1}$, $R_{2,t-2}$, $R_{1,t}$, $R_{1,t-1}$, and $R_{1,t-2}$. (v) The *p*-value of Hansen's overidentifying restrictions test results are shown in the last column. (vi) The table only reports the results that reached convergence in less than 1000 iterations and was not rejected by Hansen's (1982) test. (vii) The first column shows the method by which the weighting matrix was reached, ASI denotes Andrews (1991), NWFSI denotes Newey & West (1987) with fixed windows, and NWVSI denotes Newey & West (1994), with variable windows. (viii) The parameters γ and ψ in the Kreps–Porteus utility function model were estimated indirectly using the delta method (see Greene, 2008).

Table 4. Results of parameters β , γ , and ψ in CCAPM studies.

	β	γ	$\psi = 1/\gamma$
<i>CRRA</i>			
This paper (with Rule of Thumb)	0.9783	-0.0974	-10.2669
Issler & Piqueira (2000)	0.99	0.62	1.61
Catalão & Yoshino (2006)	0.9711	2.1192	0.47
<i>External Habits</i>			
This paper (with Rule of Thumb)	0.9793	0.0548	18.2482
Issler & Piqueira (2000)	0.99	0.46	2.17
<i>Kreps-Porteus</i>			
This paper (with Rule of Thumb)	0.9743	1.00013	3.5798
Issler & Piqueira (2000)	0.96	0.68	0.29
Bonomo & Domingues (2002)	0.9505	3.23	0.37

Tabela 5 compares the results for the rule of thumb parameter with the previous findings with Brazilian data. The estimates for the CRRA and external habits were close to the findings of Reis et al. (1998), Issler & Piqueira (2000) and Gomes (2004), $\hat{\lambda} \approx 0.80$, $\hat{\lambda} \approx 0.74$ and $\hat{\lambda} \approx 0.85$ respectively. For the Kreps-Porteus model, the median estimate was $\hat{\lambda} = 0.2263$, but in almost all cases λ estimates are not significant at the 5% level. This number differs from the results for the two preference forms treated above and for the previous studies. If we took a different approach and considered only the results significant at the 5% level, the estimate for this parameter would be $\hat{\lambda} = 0.7692$, very close to all the other results.

For the CRRA and the external habits models, which yielded various significant results for λ at the 5% level, we built confidence intervals. For the CRRA, [0.87, 0.91] and for External habits, [0.72, 0.85].

One possible reason pointed out for a high λ was the lack of credit available to the Brazilian population during the period of study. After the end of the hyperinflation in 1994, the Brazilian economy experienced a strong expansion of credit, so some agents who followed the rule of thumb due to credit constraint in the previous studies could have started to optimize their consumption decisions.

Gomes & Paz (2004) and Arreaza (2000) results for λ suggests rejection of the PIH for Latin American data. This study comes to the same conclusion for Brazilian data, but reached to slightly different values for the fraction of myopic consumption.

Table 5. Results in the literature for the estimation of consumers' share who follow the rule of thumb in the Brazilian economy.

Authors	Period studied	Estimates
Cavalcanti (1993)	1980 a 1989	0.32
Reis et al. (1998)	1947 a 1994	0.80
Issler & Piqueira (2000)	1947 a 1994	0.74
Gomes (2004)	1947 a 1999	0.85
Gomes & Paz (2004)	1951 a 2000	0.61
Gomes & Paz (2010)	1950 a 2003	[0.83, 0.91] (IPA) [0.73, 1.06] (IGP-DI)
This paper	1995 a 2011	[0.87, 0.91] (CRRA)
	1995 a 2011	[0.72, 0.85] (External Habits)



5. CONCLUSION

This paper investigated whether there is a fraction of consumers that do not behave as fully forward-looking optimal consumers in the Brazilian economy. We used different utility functional forms in the CCAPM framework. Beginning from Euler equations of the optimizing consumer utility problem, we estimated the structural parameters using the generalized method of moments (GMM) and tested the model's over-identifying restrictions using Hansen's TJ test (Hansen & Singleton, 1982).

Regarding the model's performance, we conclude that in the Brazilian case there is a proportion of the individuals consuming their current income, and another group of individuals that consume optimally in each period. These findings suggest that a significant fraction of the Brazilian disposable income went to households who consumed their current income, following the rule of thumb. Therefore, there was a strong violation of the permanent income hypothesis.

The results found can be summarized as follows:

1. The main results show that for the CRRA and external habits utilities, most of the estimates of the rule-of-thumb parameters were statistically significant at conventional levels. The interval of confidence estimates results were [0.72, 0.85] and [0.87, 0.91] respectively.
2. For the Kreps–Porteus utility function almost all estimates of λ were statistically insignificant, therefore the 22% median estimate is not robust enough to say there was a fraction of myopic consumers.
3. The results for the intertemporal discount coefficient, β , were all significant at the 5% level, for the three functional forms and in line with the previous studies.
4. Comparing the relative risk aversion results to the previous studies, we obtained lower values for the CRRA and the external habits utility models.

There are two possible explanations of the higher λ reached in the present study. One possible reason for a high λ was the lack of credit available to the Brazilian population during the period of study. After the end of the hyperinflation in 1994, the Brazilian economy experienced strong expansion of the credit, so some agents who followed the rule of thumb due to credit constraint in the previous studies could have started to optimize their consumption decisions. On the other hand, the long period with no funds to finance consumption caused a large pent-up demand during the period of this study. Another explanation is that great increase in income experienced by the lower social classes, especially after 2002, caused them to increase spending in a Keynesian way, assuming that those social classes spend their current income.

Interesting extensions of this paper could be to use factor model analysis to build portfolios in order to consider more than two assets or to explore other functional forms.

REFERENCES

- Abel, A. (1990). Asset prices under habit formation and catching up with the joneses. *American Economic Review*, 80(2), 38–42.
- Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3), 817–858.
- Arreaza, A. (2000, October 13). *Liquidity constraints and excess sensitivity of consumption in Latin American countries*. Retrieved from http://www.cemla.org/red/papers2000/v_red_arreaza.pdf
- Bonomo, M., & Domingues, G. (2002). Os puzzles invertidos no mercado brasileiro de ativos. In M. Bonomo (Ed.), *Finanças aplicadas ao Brasil* (pp. 105–120). Rio de Janeiro: FGV Editora.
- Breedon, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3), 256–296.

- Campbell, J. Y., & Mankiw, N. G. (1989, April). *Consumption, income, and interest rates: Reinterpreting the time series evidence* (NBER Working Paper No. 2924). National Bureau of Economic Research. doi: 10.3386/w2924
- Campbell, J. Y., & Mankiw, N. G. (1990). Permanent income, current income, and consumption. *Journal of Business & Economic Statistics*, 8(3), 265–79.
- Catalão, A. B., & Yoshino, J. A. (2006). Fator de desconto estocástico no mercado acionário brasileiro. *Estudos Econômicos*, 36(3), 435–463.
- Cavalcanti, C. B. (1993). Intertemporal substitution in consumption: An empirical investigation for Brazil. *Brazilian Review of Econometrics*, 13(2), 203–229.
- Cushing, M. J. (1992). Liquidity constraints and aggregate consumption behavior. *Economic Inquiry*, 30(1), 134–153. doi: 10.1111/j.1465-7295.1992.tb01540.x
- Driscoll, J., & Kraay, A. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 80(4), 549–560.
- Epstein, L. G., & Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4), 937–969.
- Epstein, L. G., & Zin, S. E. (1991). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, 99(2), 263–286.
- Flavin, M. A. (1981). The adjustment of consumption to changing expectations about future income. *Journal of Political Economy*, 89(5), 974–1009.
- Friedman, M. (1957). *A theory of consumption function*. Princeton, NJ: Princeton University Press.
- Gomes, F. A. R. (2004). Consumo no Brasil: Teoria da renda permanente, formação de hábito e restrição à liquidez. *Revista Brasileira de Economia*, 58(3), 381–402.
- Gomes, F. A. R., & Paz, L. S. (2004). Especificações para a função consumo: Testes para países da América do Sul. *Pesquisa e Planejamento Econômico*, 34(1), 39–55.
- Gomes, F. A. R., & Paz, L. S. (2010). Consumption in South America: Myopia or liquidity constraints. *Economia Aplicada*, 14(2), 129–145. doi: 10.1590/S1413-80502010000200001
- Greene, W. H. (2008). *Econometric analysis* (6th ed.). Prentice Hall.
- Hall, R. E. (1978). Stochastic implications of the Life Cycle-Permanent Income Hypothesis: Theory and evidence. *Journal of Political Economy*, 86(6), 971–87.
- Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica*, 50(5), 1269–1286. Retrieved from <http://www.jstor.org/stable/1911873>
- Hansen, L. P., & Singleton, K. J. (1983). Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy*, 91(2), 249–265. Retrieved from <http://www.jstor.org/stable/1832056>
- Issler, J. V., & Piqueira, N. S. (2000). Estimating relative risk aversion, the discount rate, and the intertemporal elasticity of substitution in consumption for Brazil using three types of utility function. *Brazilian Economic Review of Econometrics*, 20(2), 201–239.
- Kreps, D. M., & Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46(1), 185–200.
- Lucas, R. E., Jr. (1978). Asset prices in an exchange economy. *Econometrica*, 46(6), 1429–45.
- Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–08.
- Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, 61(4), 631–654.
- Reis, E., Issler, J. V., Blanco, F., & de Carvalho, L. M. (1998). Renda permanente e poupança precaucional: Evidências empíricas para o Brasil no passado recente. *Pesquisa e Planejamento Econômico*, 28(2), 233–272.
- Weber, C. (2002). Intertemporal non-separability and “rule of thumb” consumption. *Journal of Monetary Economics*, 49(2), 293–308.



A OMC Continua Promovendo o Comércio de Forma Desigual: Novas Evidências a Partir dos Anos 1990

Magnus dos Reis^{*}, Eduardo Trapp Santarossa[†], André Filipe Zago de Azevedo[‡], Sabino da Silva Pôrto Júnior[§]

Sumário: 1. Introdução; 2. A Organização Mundial de Comércio (OMC); 3. Especificação Econométrica e os Dados Utilizados; 4. Resultados Empíricos; 5. Conclusões.

Palavras-chave: Comércio Internacional; Modelo Gravitacional; Dados em Painel.

Códigos JEL: F1, C23, C21.

O artigo fornece evidências de que a OMC continuou tendo um impacto profundo sobre o comércio de seus países membros, no período 1990–2009, seguindo a contribuição original de Subramanian & Wei (2007). Para isto, utilizou-se o modelo gravitacional, através de efeitos fixos e com o estimador da PMVP, comparando a magnitude da abertura multilateral, entre países desenvolvidos e em desenvolvimento. Os resultados sugerem que os países em desenvolvimento, no período examinado, foram os que mais se beneficiaram do aumento do comércio mundial promovido pela OMC, ao contrário do que havia ocorrido, entre 1950 e 2000, conforme Subramanian & Wei (2007).

The paper provides evidences that the WTO continues to have a profound impact on its member's trade, in the period 1990–2009, as originally prompted by Subramanian & Wei (2007). Using a gravity model, based on panel data and the Poisson Pseudo Maximum Likelihood estimator, the paper compares the amount of multilateral trade liberalization in developed and developing countries. The results suggest that developing countries, during the period examined, benefited most from the increase in world trade promoted by WTO, inverting the situation identified by Subramanian & Wei (2007) in the period 1950–2000.

^{*}Doutorando em Economia Aplicada pelo Programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande do Sul-PPGE/UFRGS. E-mail: magnusdosreis@hotmail.com

[†]Mestre em Economia pela UNISINOS. E-mail: etsantar@hotmail.com

[‡]Professor do PPGE da UNISINOS e pesquisador do CNPq. E-mail: aazevedo@unisinos.br

[§]Professor do PPGE da UFRGS E-mail: sabino@ppge.ufrgs.br



1. INTRODUÇÃO

Muito embora poucos países efetivamente se aproximem de um livre comércio pleno, pode-se dizer que há um “consenso” entre os economistas de que o comércio internacional pode proporcionar uma melhor alocação de recursos e, conseqüentemente, elevar o bem-estar dos países envolvidos.¹ Com o intuito de usufruir tais benefícios, vinte e três países estabeleceram, em 1947, o *General Agreement on Tariffs and Trade* (GATT), cujo objetivo principal era promover o comércio internacional.² Desde então, tem se observado um crescimento das exportações mundiais sempre superior ao do crescimento do PIB mundial.

Apesar do crescimento expressivo do comércio internacional desde a constituição da GATT/OMC, Rose (2004) constatou que esse aumento não poderia ser atribuído à OMC. Os seus resultados sugerem que não haveria evidências suficientes de que a instituição, de fato, teria promovido o comércio internacional. No entanto, este resultado chamou a atenção de muitos economistas e Subramanian & Wei (2007), mais tarde, forneceram evidências robustas de que a OMC teria tido um forte impacto positivo sobre o comércio internacional, entre 1950 e 2000. A maior contribuição desses autores foi separar o grupo de países da OMC entre aqueles que efetivamente liberalizaram suas economias (países desenvolvidos) e os demais (composto em sua maioria por países em desenvolvimento) que, embora fizessem parte do acordo, não liberalizaram o comércio. Assim, a expansão do comércio teria sido desigual entre os países membros, conforme salientaram os autores, com o aumento do comércio se concentrando nos países desenvolvidos. Esses autores utilizaram o modelo gravitacional log linearizado com dados em painel e efeitos fixos como instrumento de mensuração.

No entanto, Santos Silva & Tenreyro (2006) fizeram duras críticas a esta abordagem econométrica, entre as quais se destacam o viés da transformação logarítmica, o fracasso da hipótese de homocedasticidade e a forma como os valores nulos são tratados. Segundo eles, esses problemas, normalmente, resultam em estimativas viesadas e ineficientes na presença de heterocedasticidade. Além disso, quando a amostra de países é grande, provavelmente haverá várias observações nas quais o valor do comércio bilateral é zero e isso implicaria em um viés de transformação logarítmica. Visando corrigir essas críticas, Santos Silva & Tenreyro (2006) sugerem estimar a equação gravitacional através da Pseudo Máxima Verossimilhança de Poisson (PMVP).

A partir dessas críticas, este artigo tem como objetivo mensurar se a OMC promoveu o comércio internacional, mantendo a separação de países proposta por Subramanian & Wei (2007), abrangendo um período mais recente, entre 1990 e 2009, quando houve uma liberalização mais intensa dos países em desenvolvimento. Além disso, seguindo a recomendação de Santos Silva & Tenreyro (2006), compara-se a magnitude da abertura multilateral, utilizando-se o modelo gravitacional com efeitos fixos e com o estimador da PMVP. Contudo, deve-se ter cuidado ao comparar diretamente os resultados obtidos neste artigo com aqueles encontrados por Subramanian & Wei (2007). De fato, esse não é o objetivo desse artigo, pois os trabalhos utilizaram estimadores distintos e diferentes períodos e países na amostra. Ainda, é provável que as estimativas encontradas por Subramanian & Wei (2007) sejam viesadas e ineficientes, uma vez que utilizaram o estimador de Efeitos Fixos e não a PMVP.

O artigo está dividido em quatro seções, além desta introdução. A seção 2 apresenta os princípios básicos, as rodadas de negociação e as principais assimetrias do sistema multilateral de comércio, conduzido pela OMC, bem como a evolução recente do regionalismo. A seção 3 faz uma breve revisão dos modelos gravitacionais, especifica a equação a ser estimada e descreve os dados utilizados. A seção 4 apresenta os resultados encontrados e, por fim, a seção 5 oferece as conclusões.

¹Dos argumentos contrários ao livre comércio, os dois mais refinados e aceitos são: melhoria nos termos de troca, válida apenas para a imposição de tarifas de importação em países grandes e subsídios a setores intensivos em tecnologia, dadas as externalidades positivas que este setor pode gerar.

²Em 1995 o acordo foi incorporado à Organização Mundial de Comércio (OMC).

2. A ORGANIZAÇÃO MUNDIAL DE COMÉRCIO (OMC)

A OMC é a instituição responsável por promover as rodadas de negociações multilaterais (multilateralismo) e regulamentar os Acordos Preferenciais de Comércio (Regionalismo). Além de regular os princípios básicos de todo sistema, ela também é a instituição responsável por solucionar eventuais controvérsias.³ A seguir, são apresentadas algumas características da abertura multilateral e regional.

2.1. O Multilateralismo

As características deste sistema envolvem um comércio sem discriminação, o mais previsível possível, mais competitivo e com mais benefícios aos países menos desenvolvidos. A seguir são apresentados os princípios da instituição, um pequeno histórico das rodadas de negociações multilaterais e as assimetrias do sistema.

2.1.1. Princípios do sistema de comércio

O principal princípio do sistema de comércio multilateral conduzido pela OMC é a Cláusula da Nação Mais Favorecida (MFN). Ela estabelece que os seus países membros devem conceder tratamento igual a todos os demais países da instituição, isto é, os países não poderão favorecer apenas um único país, através de uma redução tarifária, por exemplo, sem que isso seja estendido a todos demais membros da OMC. Outro princípio importante se refere a questão do Tratamento Nacional, que exige que produtos e serviços estrangeiros sejam tratados da mesma forma que são tratados os produtos e serviços domésticos, após as mercadorias estrangeiras entrarem no mercado doméstico. O mesmo se aplica para marcas estrangeiras, direitos autorais e patentes.⁴

Outro aspecto é a luta por um comércio gradualmente mais livre através das negociações comerciais. Muito embora a redução das barreiras comerciais seja um dos meios mais óbvios de incentivar o comércio,⁵ outras questões começaram a serem debatidas, tais como aspectos burocráticos, barreiras não tarifárias e políticas cambiais. No entanto, muitas vezes a abertura comercial vai requerer ajustes internos e, devido a isso, os acordos da OMC permitem que os países introduzam essas mudanças de maneira gradual, através de “liberalização progressiva”. Além disso, países em desenvolvimento recebem mais tempo para cumprir as suas obrigações (OMC, 2013).

No que diz respeito à previsibilidade, os países devem ser o mais transparente possível e suas políticas comerciais devem ser críveis para reduzir as incertezas. A promessa de não aumentar uma barreira comercial pode ser tão importante como a redução de uma, porque ela dá às empresas uma visão mais clara de suas oportunidades futuras. Com estabilidade e previsibilidade, o investimento é estimulado, empregos são criados e os consumidores podem desfrutar plenamente dos benefícios de uma maior concorrência, escolha e preços mais baixos. Portanto o sistema multilateral de comércio é uma tentativa, por parte dos governos, de tornar o ambiente de negócios estável e previsível (OMC, 2013).

Outro aspecto interessante é impulsionar uma concorrência leal. Muitas vezes a OMC é definida como uma instituição de “livre comércio”, no entanto isto não é inteiramente verdadeiro, pois em circunstâncias limitadas o sistema permite a imposição de tarifas de importação e outras formas de proteção. Na

³O mecanismo de resolução de disputas comerciais envolve cinco etapas, desde consultas iniciais entre os países diretamente envolvidos até a implementação de um painel de especialistas. A duração varia de 1 ano (sem apelação) a 15 meses (com apelação) e ela dá o direito de retaliação (OMC–World Trade Organization, 2013).

⁴Dado que o tratamento nacional aplica-se apenas após os produtos ingressarem no mercado doméstico, a cobrança de tarifa de importação não é considerada uma violação do tratamento nacional, mesmo se os produtos produzidos localmente não sejam taxados com um imposto equivalente. Um exemplo de violação desta cláusula seria cobrar um imposto nacional (que abrange todos os produtos do mercado independentemente de qual for sua procedência) mais alto de produtos importados do que os nacionais (OMC, 2013).

⁵Entendem-se como barreiras os direitos aduaneiros (ou tarifas), medidas como a proibição de importação e cotas que restringem quantidades importadas.



realidade, o sistema pode ser mais bem definido como um sistema de regras dedicadas à concorrência aberta, justa e leal. As regras da MFN e do Tratamento Nacional são projetadas para garantir condições “justas” de comércio. Porém, muitos países praticam o dumping e concedem subsídios, práticas essas que são consideradas desleais pela OMC. Para esses casos, a instituição permite que os governos possam se defender através da cobrança de direitos de importação adicionais para compensar os danos causados por comércio desleal (OMC, 2013). Por fim, a instituição incentiva o desenvolvimento e a reforma econômica, pois países em desenvolvimento necessitam de maior flexibilidade temporal para implementarem acordos e este benefício é concedido apenas aos mesmos (OMC, 2013).

Apesar das limitações ao livre comércio permitidas pela própria OMC, o período de maior crescimento do volume de comércio internacional no século XX ocorreu justamente após a criação do GATT, como mostra a Tabela 1. No período 1950–1973, o crescimento médio anual do volume de comércio atingiu a 8,2%, muito acima do período entre guerras (1913–1950) e superior também ao início do século, quando se vivia o auge da chamada 1ª onda de globalização. Mesmo após o término do acordo de Bretton-Woods, em 1973, o comércio manteve taxas de crescimento significativas, superiores ao da primeira metade do século XX. Embora o aumento do comércio internacional não possa ser atribuído exclusivamente à formação da OMC, uma parcela importante teve a contribuição deste organismo. De acordo com Subramanian & Wei (2007), a OMC seria responsável por um comércio internacional adicional da ordem de 120% (ou US\$ 8 trilhões), apenas no ano de 2000.

2.1.2. As rodadas de negociações multilaterais

Desde 1947, já foram realizadas 8 rodadas de negociações multilaterais e, atualmente, a rodada de Doha está em curso. Conforme pode ser observado na Tabela 2, a duração, o número de países membros e a complexidade dos temas foram aumentando a cada rodada. Nas rodadas iniciais, o foco era a negociação do comércio de produtos industrializados. Especificamente nas seis primeiras rodadas, o tema predominante foi a busca de reduções tarifárias. A partir da sétima rodada, diversos novos temas foram incluídos nas negociações, incluindo medidas antidumping, subsídios, salvaguardas e compras governamentais. Devido ao maior número de países-membros, que já chegava a 99 na época de sua realização, e a maior complexidade dos temas, a rodada Tóquio se estendeu por seis anos, a mais longa até então.

Na rodada Uruguai outras temas foram acrescentados na pauta de negociação, tais como a preservação das patentes sobre propriedade intelectual (TRIPS) e a liberalização do setor de serviços. Esses novos temas têm grande importância nos mercados dos países em desenvolvimento, que até então estavam praticamente excluídos do processo de negociação. Também houve significativos avanços nas negociações sobre as regras da OMC em matéria de subsídios, barreiras técnicas, salvaguardas, regras de origem, licenças de importação, antidumping, medidas fitossanitárias, e investimentos relacionados ao comércio (TRIMS), entre outros. Além disso, para tornar o comércio mais previsível, onde os agentes econômicos tivessem confiança de que as barreiras não seriam elevadas arbitrariamente, houve um aumento expressivo na imposição de tetos para tarifas e barreiras não-tarifárias para os países em desenvolvimento. No caso de produtos industrializados, o percentual de produtos com tetos tarifários passou de 21% para 73% nesse grupo de países (GATT, 1994).

No que se refere ao valor das tarifas, houve a redução das tarifas de importação de bens industriais em países desenvolvidos de 40% (de uma média de 6,3% para 3,8%). No caso dos países em desenvolvimento,

Tabela 1. Crescimento do Volume de Comércio e PIB Mundial (1900–2010). *

Variação média anual	1900–1913	1913–1950	1950–1973	1973–2000	2000–2010
PIB	2,5	2,0	5,1	2,9	3,0
Comércio de Bens	4,3	0,6	8,2	4,9	5,0

* Variação percentual média ao ano. Fonte: OMC.

Tabela 2. Rodadas de negociações multilaterais realizadas.

Ano	Local	Nome	Nº de países	Temas
1947	Genebra		23	tarifas
1949	Annecy		13	tarifas
1951	Torquay		38	tarifas
1956	Genebra		26	tarifas
1960–1961	Genebra	Rodada Dillon	26	tarifas
1964–1967	Genebra	Rodada Kennedy	62	tarifas e medidas antidumping
1973–1979	Genebra	Rodada Tóquio	102	tarifas, barreiras não tarifárias e estrutura institucional
1986-1994	Genebra	Rodada Uruguai	123	tarifas, barreiras não tarifárias, serviços, propriedade intelectual, solução de controvérsias, têxteis, agricultura, criação da OMC
2001–?	Doha	Rodada Doha	155	mesmos temas acima, com ênfase na agricultura e têxteis

Fonte: Elaboração própria com base no OMC (2013).

a redução foi variável, mas bastante significativa para a maioria dos países. No caso de Índia, Coreia do Sul e Cingapura, a tarifa média de importação declinou em mais de 50% para produtos industrializados. Na Índia, a queda foi maior, passando de uma tarifa média de 71,4% antes da rodada Uruguai para 32,4% após a rodada (GATT, 1994).⁶ De acordo com Thorstensen (2001), todas as oito rodadas da OMC já concluídas podem ser consideradas um sucesso, quando se tem em mente que as médias tarifárias aplicadas aos produtos eram de 40% em 1947 e caíram para 5% em 1994, quando se finalizou a Rodada Uruguai.

A rodada de Doha, lançada em 2001, tinha objetivos ambiciosos, pois abrangia não apenas o liberalizar o comércio de manufaturas, produtos agrícolas e serviços, mas também uma série de temas relacionados indiretamente com o comércio (tais como propriedade intelectual, defesa da concorrência estrangeira, regras de investimento, entre outros). A meta inicial era concluir as negociações para a grande parte desses temas até 1º de Janeiro de 2005. No entanto, essa rodada permanece aberta há mais de 11 anos e as negociações seguem sem avançar, o que põe em cheque o futuro do sistema.

Isso reflete dois problemas estruturais na forma como são feitas as negociações na OMC. O primeiro se deve a dois fatos: que as decisões na OMC devem ser consensuais e ao aumento do número de países membros. Uma vez que as negociações comerciais iniciaram, em 1947, com apenas 23 países e que, quando Doha foi lançada eram 155, a abertura multilateral não está conseguindo avançar devido à dificuldade de consenso sobre os temas. Já o segundo é decorrente do audacioso objetivo inicial, uma vez que a ideia era conseguir um grande acordo no qual englobaria manufaturas, serviços e produtos agrícolas. No entanto, não foi possível chegar a um acordo para o setor agrícola, pois os países desenvolvidos não abriram mão de suas políticas protecionistas nesse setor. Esta situação vem frustrando muitos países, fazendo com que o multilateralismo fosse preterido ao regionalismo no mundo inteiro (The Economist, 2012).

⁶Outros aspectos importantes negociados na rodada Uruguai foram o estabelecimento de regras mais precisas para medidas antidumping e direitos compensatórios e a eliminação das “Restrições Voluntárias às Exportações” até 1999.



2.1.3. Assimetrias da OMC, contradições e evidências empíricas

A primeira das assimetrias existentes na OMC, segundo Subramanian & Wei (2007), é a diferença de tratamento que a OMC fez entre países desenvolvidos e em desenvolvimento. Os membros desenvolvidos tiveram um cronograma mais rigoroso de obrigações do que os países em desenvolvimento, em termos de liberalização comercial (barreiras tarifárias e não tarifárias). A segunda assimetria está relacionada à discriminação feita por países membros a não membros da OMC, uma vez que as obrigações de reciprocidade e o status de MFN eram estendidos apenas a membros da instituição. A terceira se deve a diferença na forma que ocorreu a liberalização comercial entre os setores. Como os países desenvolvidos foram os precursores da instituição, foram eles que participaram mais ativamente das rodadas de negociações iniciais e, assim, era natural que fossem negociados primeiramente os setores nos quais eles tinham interesse em exportar, enquanto os setores nos quais os países em desenvolvimento tinham interesse ficaram à margem das negociações. Por fim, a quarta assimetria reflete a diferença de tratamento que os novos países membros recebem em comparação aos que ingressaram antigamente. Os novos membros recebem um cronograma de liberalização mais rigoroso se comparado aos países que são membros antigos da OMC.

Levando em conta essas assimetrias e utilizando o modelo gravitacional, Subramanian & Wei (2007) contrariam os resultados Rose (2004) ao fornecerem evidências robustas de que a OMC tem tido um forte impacto positivo no comércio, gerando um impacto adicional de cerca de 120% no comércio mundial. No entanto, o impacto entre os países teria sido desigual. Primeiro, os países industriais, que participaram mais ativamente do que os países em desenvolvimento nas negociações comerciais recíprocas, beneficiaram-se de um grande aumento do comércio. Em segundo lugar, o comércio bilateral foi maior quando ambos os parceiros se comprometeram com a liberalização, quando comparado a uma situação em que apenas um parceiro efetivamente liberalizou sua economia. Terceiro, os setores que não se beneficiaram do processo de liberalização não usufruíram de aumento no comércio bilateral.

Essas assimetrias, de fato, criaram durante muito tempo um viés de comércio entre os membros da OMC, que realmente reduziram o grau de protecionismo para os demais integrantes do acordo, abrangendo principalmente os países desenvolvidos. No entanto, em um período mais recente, especialmente a partir da rodada Uruguai, os países em desenvolvimento passaram a efetivamente liberalizar o comércio, seguindo as orientações da OMC. O exemplo mais claro disso foi a imposição de tetos tarifários, que reduziram significativamente as tarifas de importação até então vigentes para os demais membros da OMC. Ao mesmo tempo, a liberalização dos países desenvolvidos foi mais tímida, principalmente porque já eram economias muito mais abertas ao comércio, por terem reduzido seu aparato protecionista no período pré-rodada Uruguai (1947–1986). Ou seja, a partir da metade dos anos 1990, a liberalização mais intensa ocorreu nos países em desenvolvimento que eram membros da OMC.

O resultado dessa maior liberalização pode ser observado pelo aumento nos fluxos de comércio dos países em desenvolvimento pertencentes à OMC. Conforme mostra a Tabela 3, no período 1990–2009, as importações desse grupo de países cresceram, em média, 10,5% ao ano. Já os países desenvolvidos tiveram um crescimento menor de suas importações, no mesmo período, chegando a 5,8% ao ano. Portanto, a principal assimetria de comércio entre os membros da OMC, originalmente destacada por Subrama-

Tabela 3. Evolução das Importações dos Países da OMC por grau de Desenvolvimento. *

Todos países da OMC	1990 (US\$ milhões)	2009 (US\$ milhões)	Média de Crescimento 1990–2009
Países desenvolvidos importadores	2.584.353,92	7.561.283,93	5,81%
Países em desenvolvimento importadores	637.605,76	4.278.625,09	10,54%

* Classificação de acordo com a UNCTAD. Fonte: OMC.

nian & Wei (2007), parece ter se invertido a partir dos anos 1990. No período mais recente, especialmente com as medidas adotadas a partir da Rodada Uruguai, foram os países em desenvolvimento que mais liberalizaram suas economias de forma multilateral e, como consequência disso, foram eles que experimentaram o maior incremento comercial.

Por fim, outro aspecto constantemente debatido e criticado no sistema de comércio é que existem algumas brechas que tornam possível violar os princípios básicos da OMC. A violação mais comum é a formação de APCs que podem liberalizar o comércio apenas dentro do bloco, muitas vezes discriminando as mercadorias extrabloco.⁷ Assim, os blocos econômicos contrariam a cláusula MFN. É possível também que países desenvolvidos concedam a países em desenvolvimento um acesso preferencial aos seus mercados sem estender este mesmo benefício a países desenvolvidos. Esta brecha é conhecida como Sistema Geral de Preferência (GSP) e também viola a cláusula da MFN. Por fim, um país também pode elevar suas barreiras tarifárias contra produtos de países que são negociadas de “forma injusta” dentro do país lesado, este é o caso do *dumping*.⁸

2.2. Regionalismo

O regionalismo não é um fenômeno recente. Segundo Bhagwati, Krishna & Panagariya (1999), a primeira onda do regionalismo ocorreu no século XX, mais especificamente no final da década de 1950 e durante a década de 1960. Essa primeira onda envolvia apenas setores industriais específicos, removendo unicamente as barreiras de importação, era composta somente por países com o mesmo nível de desenvolvimento e foi dirigida pelos governos, pois era baseada na política de substituição de importações. Já a segunda onda envolve quase todos os setores, inclui países com diferentes níveis de desenvolvimento, disciplinas comuns, regimes regulatórios e tem sido liderada pelas empresas ao buscarem novos mercados para suas exportações. Nesse sentido, a segunda onda do regionalismo, que ocorre a partir de 1990, tem uma integração mais profunda em comparação à primeira. Além disso, houve um aumento expressivo na criação de acordos bilaterais e de blocos econômicos no mundo inteiro. A Figura 1 demonstra justamente este crescimento no número de APCs registrados na OMC. Para se ter uma ideia, o número total de APCs em vigor notificados à OMC em 1999 era maior que o número de países membros desta instituição.

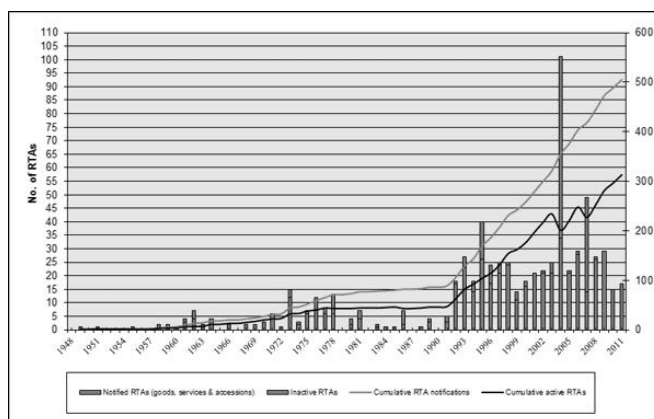
Após a apresentação das características do sistema de comércio, a próxima a seção apresenta o modelo gravitacional, que foi o instrumento utilizado para descobrir se a OMC conseguiu promover o comércio mundial e para comparar a abertura multilateral à regional. A escolha pela equação gravitacional se deve ao fato de que ela tem sido o principal instrumento utilizado pelos pesquisadores para mensurar o impacto dos APCs, uma vez que ela é capaz de explicar o volume de comércio bilateral de forma muito satisfatória mesmo se os APCs não tivessem sido formados.

3. ESPECIFICAÇÃO ECONÔMETRICA E OS DADOS UTILIZADOS

O modelo gravitacional tem origem na lei da gravitação universal de Isaac Newton, formulada na metade do século XVII, na qual postula que a força de atração entre dois corpos é diretamente proporcional ao produto de suas massas e inversamente proporcional ao quadrado de sua distância. No século XIX, a lei gravitacional foi absorvida por diversas áreas do conhecimento humano, tais como as ciências sociais e exatas. Mais especificamente na ciência econômica, Tinbergen (1962), Poyhonen (1963) e Linnemann (1966) adaptaram a lei da gravidade à economia para estimar o fluxo de comércio bilateral entre dois

⁷ O artigo XXIV é o que regula os Acordos Preferenciais entre países desenvolvidos e a *Enabling Clause*, a que estipula as regras para Acordos Preferenciais entre países em desenvolvimento.

⁸ Muito embora a OMC só permita essas exceções ao livre comércio sob condições estritas, em geral, a cláusula da MFN garante que quando um país reduz uma barreira comercial, ou ainda, quando abre seu mercado, isso deverá ser feito para todos os seus parceiros comerciais, sejam eles países ricos ou pobres.

**Figura 1.** Número de notificações de APCs à OMC.

Fonte: OMC (2013).

países. Foram esses autores que propuseram o modelo gravitacional ao fornecer as variáveis básicas que são utilizadas, em alguns casos, até hoje para determinar o comércio bilateral entre os países.

A equação gravitacional tradicional parte da ideia básica de que o comércio entre dois países é diretamente proporcional ao Produto Interno Bruto (PIB) de suas economias e inversamente proporcional à distância entre eles. Sua forma mais simples pode ser descrita da seguinte maneira:

$$\ln m_{ji} = \beta_0 + \beta_1 \ln Y_j + \beta_2 \ln \left(\frac{Y_j}{N_j} \right) + \beta_3 \ln Y_i + \beta_4 \ln \left(\frac{Y_i}{N_i} \right) + \beta_5 \ln \text{dist}_{ji} + \epsilon_{ji}, \quad (1)$$

onde m_{ji} representa as importações bilaterais entre os países j e i ; Y_i é o PIB do país i e Y_j é o PIB do país j ; N_i é a população do país i e N_j é a população país j ; dist_{ji} é a distância entre os países j e i ; e ϵ_{ji} é o erro.

Entretanto, apesar de conseguir explicar, de forma satisfatória, o comércio bilateral entre os países, o modelo gravitacional carecia de fundamentação teórica. Assim, Anderson (1979), Deardorff (1998), e Anderson & van Wincoop (2003) derivaram o modelo gravitacional do modelo de Heckscher-Ohlin, assim como Bergstrand (1985), Helpman (1987) e Helpman & Krugman (1995) também, porém através dos modelos de concorrência imperfeita, dando a base teórica que faltava ao modelo. Mais recentemente, Baldwin & Taglioni (2006) também fundamentaram a equação gravitacional.

Além disso, com o passar dos anos, os pesquisadores perceberam que outros fatores poderiam afetar o volume de comércio bilateral e, assim, passaram a incluir na equação gravitacional outras variáveis, a fim de aumentar o poder de explicação do modelo. Nesse sentido, *dummies* foram inseridas para representar algumas características que se presumia que também afetassem o volume de comércio entre os países. Por exemplo, espera-se que países com um idioma oficial em comum e que partilhem uma fronteira tenham maiores facilidades de negociar e, sendo assim, apresentem um volume de comércio maior em comparação aos países que não apresentem tais características. Além disso, Aitken (1973) foi um dos pioneiros a avaliar o impacto da criação dos APCs sobre os fluxos de comércio. Nesse caso, o modelo gravitacional refletia o comércio “normal” entre dois países sem a presença de um APC, enquanto a variável *dummy* relacionada ao APC captava o comércio que podia ser atribuído exclusivamente à criação do mesmo.

Nesse primeiro momento, a maior parte dos trabalhos realizados estimou o modelo gravitacional através de Mínimos Quadrados Ordinários (MQO) e com dados em *cross-section* ou dados agrupados (*pooled data*). Esta arquitetura pode ser considerada a abordagem tradicional do modelo. Entretanto, as estimativas tradicionais dos modelos gravitacionais passaram a ser alvo novamente de críticas. Wei (1996),

Deardorff (1998), e Anderson & van Wincoop (2003) afirmaram que o modelo gravitacional usual pode apresentar problemas de especificação ao ignorar a “resistência multilateral” e “isolamento” dos países. Além disso, segundo Egger (2005), o modelo gravitacional com dados agrupados e em *cross-section* sofre problemas de especificação, uma vez que ele não é capaz de lidar com a heterogeneidade bilateral do exportador e importador, que é extremamente provável que esteja presente nos fluxos de comércio bilaterais. Ademais, a teoria econômica internacional avançou e novos conceitos sobre o regionalismo foram desenvolvidos. Um deles foi proposto por Wonnacott & Lutz (1989), Summers (1991) e Krugman (1991), ao criarem o conceito de “parceiro natural de comércio” para países próximos geograficamente e que possuam elevados níveis de comércio bilateral. Esse novo conceito teve um impacto direto sobre a forma que vinha sendo estimada a equação gravitacional, fazendo com que uma nova forma de estimação fosse proposta. Assim, visando corrigir todos esses problemas de má especificação, Matyas (1997), Anderson & van Wincoop (2003) e Baldwin & Taglioni (2006) sugerem a estimação da equação gravitacional em dados em painel e efeitos fixos.

Apesar de certo consenso da necessidade de estimar o modelo com dados em painel e efeitos fixos, a maneira de especificar o efeito fixo e qual estimador utilizar ainda vem provocando um grande debate.⁹ Santos Silva & Tenreyro (2006) e Magee (2008) defenderam a utilização do estimador PMVP, enquanto Cheng & Wall (2004) e Baldwin & Taglioni (2006) trabalharam com o estimador LSDV. Por outro lado, o estimador HT foi utilizado por Egger (2005) e Carrere (2006). Contudo, Santos Silva & Tenreyro (2006) fizeram duras críticas à prática usual de estimar o modelo gravitacional com transformação logarítmica. Segundo eles, essa prática incorre em um viés da transformação logarítmica, dada a forma como os valores nulos são tratados, e fracassa na hipótese de homocedasticidade. Os autores afirmam ainda que esses problemas, normalmente, resultam em estimativas viesadas e ineficientes na presença de heterocedasticidade.¹⁰

Santos Silva & Tenreyro (2006) utilizaram simulações de Monte Carlo para comparar vários estimadores para o modelo gravitacional, entre eles: MQO linear e não linear, *tobit* e a Pseudo Máxima Verossimilhança de Poisson (PMVP).¹¹ Os autores demonstraram que a melhor forma de estimar o modelo gravitacional seria através da PMVP, pois as estimativas dos parâmetros são consistentes mesmo com erros heterocedásticos e ainda é possível incluir na análise os fluxos de comércio bilaterais com valores iguais a zero sem fazer qualquer transformação. Magee (2008); Recalde, Florensa & Iturralde (2008); Siliverstovs & Schumacher (2007); Westerlund & Wilhelmsson (2009) corroboram com a visão de Santos Silva & Tenreyro (2006), ao afirmarem que os resultados encontrados pela equação gravitacional são muito sensíveis ao método utilizado e que o PMVP seria o mais confiável.

Outra alternativa para lidar com o problema de zeros na amostra é a proposta por Helpman, Melitz & Rubinstein (2008). No artigo, os autores derivam um modelo gravitacional a partir de um modelo de firmas heterogêneas e advertem para o possível viés nas estimativas dos parâmetros tanto pela existência de zeros (viés de seleção) quanto pela existência de heterogeneidade das firmas. Resumidamente, a sugestão para correção dos problemas é estimar o modelo gravitacional em dois estágios: primeiramente, através de um modelo *probit*, com o objetivo de identificar a probabilidade de o comércio ser igual a

⁹Para maiores informações sobre formas alternativas de especificar o efeito fixo veja Reis (2012).

¹⁰Segundo Santos Silva & Tenreyro (2006), a desigualdade de Jensen afirma que $E[\ln y] \neq \ln E[y]$, ou seja, que o valor esperado do logaritmo de uma variável aleatória é diferente do logaritmo do valor esperado, e isso implica que a prática padrão de interpretar os parâmetros dos modelos log linearizados estimados por MQO em termos de elasticidades pode ser altamente enganadora na presença de heterocedasticidade, uma vez que o valor esperado do logaritmo de uma variável aleatória depende dos momentos de ordem superior de sua distribuição. Portanto, se os erros são heteroscedásticos, os erros transformados serão geralmente correlacionado com as covariáveis.

¹¹Segundo Wooldridge (2010, p.558), a hipótese de normalidade como distribuição padrão da regressão linear nem sempre é a melhor alternativa, pois a distribuição normal é de variáveis contínuas que podem assumir quaisquer valores e isso não ocorre para dados de contagem. Nos dados de contagem, a variável dependente pode assumir poucos valores e sua distribuição pode ser muito diferente da distribuição normal. Nesse sentido, para dados de contagem a distribuição mais indicada seria a de Poisson. Quando se refere à Pseudo Máxima Verossimilhança de Poisson, se assume que a distribuição de Poisson não seja inteiramente correta nos dados utilizados.



zero e, em seguida, adiciona à equação final essa probabilidade. Contudo, o estimador escolhido para estimador do modelo gravitacional foi a PMVP.

Diferentemente do artigo de Subramanian & Wei (2007), que estimaram a equação log linearizada com dados de cinco em cinco anos para o período de 1950 a 2000, o presente artigo estimará a equação gravitacional através da PMVP e dois efeitos fixos (α_{ji} e α_t), aliviando qualquer risco de incorrer em vies da transformação logarítmica e produzindo resultados consistentes mesmo na presença de erros heterocedásticos. O efeito fixo α_{ji} captará todas as características não observadas, que são constantes ao longo do tempo para cada par de países, mas que influenciam os fluxos de comércio. Nesse sentido, leva-se em conta a possibilidade de existência de parceiros naturais de comércio, a resistência multilateral, o isolamento dos países, aspectos culturais e institucionais, aliviando as críticas anteriores. Todas essas características não observadas, em alguns casos, não são mensuráveis, porém elas afetam os fluxos de comércio e devem ser levadas em conta pelo modelo. Justamente dessa maneira que um possível problema com heterogeneidade é controlado no modelo. Não obstante, a inclusão do efeito fixo α_{ji} leva em conta inclusive as variáveis observáveis que usualmente são utilizadas na equação gravitacional, entre as quais a distância entre os países, a extensão da área territorial do importador e do exportador, bem como as *dummies* de fronteira, litoral e idioma em comum.¹² Por outro lado, o efeito fixo específico de cada ano t , α_t , captura choques não percebidos pelas variáveis incluídas no modelo, tais como o ciclo de negócios, fatores climáticos, guerras, entre outros, que podem determinar mudanças nos fluxos de comércio.

Além dos dois efeitos fixos, a equação a ser estimada terá como variável dependente as importações dos países e como independentes o PIB e a população do importador e exportador. Além disso, foram inseridas várias variáveis *dummies* para capturar o efeito da abertura regional e multilateral. As *dummies* incluídas no modelo foram construídas, conforme descrito na Tabela 4. Seguindo Soloaga & Winters (2001), foram incluídas três variáveis *dummies* para cada bloco, sendo uma delas para capturar o efeito da criação do APC sobre o comércio intrabloco, outra para as importações totais do bloco e a última para as exportações totais do bloco. Através da inserção dessas variáveis, esses autores conseguiram analisar não somente o impacto sobre o fluxo de comércio bilateral que pode ser exclusivamente atribuído à formação do bloco, mas também se houve criação ou desvio de comércio com a formação do mesmo, bem como se ocorreu o que ambos os autores denominam de “desvio de exportação”. Por outro lado, as *dummies* que representam a abertura multilateral promovida pela OMC foram embasadas pelo artigo de Subramanian & Wei (2007).

Assim, parte-se agora para definir o modelo a ser estimado neste artigo, representado pela equação (2). A matriz X_{jit} inclui o PIB e a população dos importadores e exportadores, além das *dummies* da abertura regional e multilateral. Portanto, os coeficientes de interesse serão estimados pelo vetor coluna β_k .

$$m_{jit} = \alpha_t + \alpha_{ji} + \sum_{k=1}^n \beta_k \cdot X_{jit} + \epsilon_{jit}, \quad (2)$$

onde m_{jit} é vetor coluna das importações dos países importadores j e exportadores i no ano t ; X_{jit} é a matriz com as k variáveis dos importadores j e exportadores i no ano t ; β_k é o vetor de parâmetros das variáveis X_{jit} ; α_{ji} é o efeito fixo das características não observadas que são constantes no tempo para os pares de países ji ; α_t é o efeito fixo específico de cada ano t que é comum para todos pares de países ji ; e ϵ_{jit} é o vetor de resíduos randômicos do ano t .

Finalmente, no que diz respeito aos dados, a amostra utilizada engloba 52 países para um período de 20 anos, de 1990 a 2009.¹³ Estas nações representavam, em termos de comércio mundial, aproximadamente, 78% das importações mundiais totais para o período analisado. O trabalho apresenta 2.652

¹² Mesmo essas variáveis não sendo colineares com o efeito fixo α_{ji} , não será possível mensurá-las, nem quaisquer outras, que sejam constantes no tempo, quando o efeito fixo for assim especificado.

¹³ A lista dos países incluídos na análise é a seguinte: Argentina, Brasil, Paraguai, Uruguai, Canadá, México, Estados Unidos, Peru, Equador, Bolívia, Colômbia, Venezuela, Cingapura, Filipinas, Indonésia, Malásia, Tailândia, Austrália, Nova Zelândia, Costa Rica,

Tabela 4. *Dummies* inseridas no modelo.

Sigla	Forma de construção
APC	<i>Dummy</i> que assume valor 1 se o importador <i>i</i> e o exportador <i>j</i> forem membros de um APC e, 0 caso contrário.
APC_OMC	<i>Dummy</i> que assume valor 1 se o importador <i>i</i> é membro de um APC, porém o exportador <i>j</i> é membro da OMC e não faz parte do acordo e 0, caso contrário.
APC_NOMC	<i>Dummy</i> que assume valor 1 se o importador <i>i</i> é membro de um APC, porém o exportador <i>j</i> não faz parte do acordo e não é membro da OMC e 0, caso contrário.
OMCD_OMCD	<i>Dummy</i> que assume valor 1 se ambos <i>i</i> e <i>j</i> (importador e exportador) são países desenvolvidos membros da OMC e não possuem um APC, e 0, caso contrário.
OMCD_OMCS	<i>Dummy</i> que assume valor 1 quando o país importador, desenvolvido, <i>i</i> e o país exportador, subdesenvolvido, <i>j</i> , ambos forem membros da OMC e sem a presença de um APC e 0, caso contrário.
OMCD_NOMC	<i>Dummy</i> que assume valor 1 quando o país importador <i>i</i> for desenvolvido e membro da OMC e o exportador <i>j</i> não seja membro da OMC e nem tenham um APC em comum e 0, caso contrário.
OMCS_OMCS	<i>Dummy</i> que assume valor 1 para quando o importador <i>i</i> e o exportador <i>j</i> forem países subdesenvolvidos, membros da OMC e que não façam parte de um APC, e 0, caso contrário.
OMCS_NOMC	<i>Dummy</i> que assume valor 1 para o país subdesenvolvido importador <i>i</i> for membro da OMC e país exportador <i>j</i> não seja membro OMC e ambos não tenham um APC em comum e 0, caso contrário.
NOMC_NOCM	<i>Dummy</i> que assume valor 1 quando o importador <i>i</i> e o exportador <i>j</i> não sejam membros da OMC e não façam parte de um APC e, 0 caso contrário.

observações anuais (52 países importadores \times 51 fluxos de importação bilateral) e 53.040 observações para todo período (2.652 observações \times 20 anos). A variável dependente, importações bilaterais em dólares americanos correntes, foi obtida a partir do UN COMTRADE. Já o PIB dos países, também medido em dólares americanos correntes, e a população foram obtidos do Banco Mundial. As *dummies* referentes aos APCs e a OMC foram construídas pelos autores. Os países foram separados entre dois grupos, os desenvolvidos e os desenvolvimentos. Os critérios utilizados foram com base nos dados do Banco Mundial.¹⁴

4. RESULTADOS EMPÍRICOS

Antes de apresentar os resultados encontrados, cabe ressaltar que o modelo foi submetido a alguns testes. O primeiro teste econométrico realizado buscava descobrir se era necessário controlar a heterogeneidade bilateral do exportador e do importador. Assim, o teste utilizado foi baseado na proposta de Greene (2008), que testa se há diferença entre os pares. Esse teste apresenta a hipótese nula (H_0) de

Egito, Islândia, Noruega, Suíça, Áustria, Finlândia, Suécia, Marrocos, Tunísia, Alemanha, Bélgica, Chipre, Dinamarca, Espanha, França, Grécia, Holanda, Hungria, Inglaterra, Irlanda, Itália, Polônia, Portugal, Romênia, Chile, China, Coreia do Sul, Hong Kong, Índia, Israel, Japão e Turquia.

¹⁴Os países desenvolvidos foram considerados países de alta renda membros da OCDE. Os países em desenvolvimento o restante. Para mais detalhes, ver <http://www.worldbank.org>.



que o efeito fixo dos pares de países (α_{ji}) é o mesmo para todos os pares. Dessa forma, não rejeitando a hipótese nula, a estrutura de dados agrupados seria a indicada. Por outro lado, quando ocorre a rejeição da hipótese nula, o teste indica a necessidade de controlar a heterogeneidade, ou seja, exige-se a utilização de dados em painel. Conforme observado na Tabela 5, o teste rejeitou fortemente a hipótese nula. Além desse, foi testada a necessidade de se incluir *dummies* de tempo para controlar o ciclo de negócios e quaisquer outros choques não percebidos. Novamente, o teste rejeitou a hipótese nula de que as *dummies* temporais são iguais a zero e, desta forma, elas também devem ser inseridas no modelo.

No que diz respeito aos resultados, eles se encontram na Tabela 6. Cabe salientar ainda que os erros são robustos para levar em conta a presença de heterocedasticidade e autocorrelação serial dos resíduos. Percebe-se que o PIB do importador e do exportador foram significativos a 99% e apresentaram o sinal de acordo com o esperado, porém o PIB do importador teve maior importância do que o do exportador. Já a população foi significativa a 99% apenas para o importador, apresentando um sinal negativo.

O coeficiente associado aos pares de países que ambos são membros da OMC, porém o importador j é membro de um APC no qual o exportador i não é membro, não foi significativo (apc_omc). Assim, não é possível afirmar que a criação da OMC incrementou o comércio fora do bloco. Se analisar essa *dummy* em conjunto com as *dummies* dos APCs, que foram significativas e positivas (Mercosul, Nafta, União Europeia, Asean e EFTA), o resultado sugere que a abertura regional foi discriminatória, uma vez que ela não foi estendida à abertura multilateral. Mais discriminatório ainda foi o comportamento associado aos importadores j , membros da OMC e de um APC no qual o exportador i não é membro do acordo e nem membro da OMC (apc_nomc), uma vez que a variável apresentou um sinal negativo e significativo a 99%. Isso significa que a não participação na OMC teve um efeito negativo nesse tipo de comércio. Observa-se então que existiu discriminação também no processo de abertura multilateral já que os países membros não foram obrigados a estenderem os mesmos privilégios da OMC aos não membros, o que, de certa forma, era de se esperar dado o principal que orienta a instituição, o princípio da MFN. Esses resultados divergem do estimado por Subramanian & Wei (2007), que encontraram impactos positivos e significativos em ambos coeficientes, devido ao fato de que alguns países terem realizado liberação tarifária unilateral, e o de Rose (2004), o qual não encontrou benefícios na adesão dos países à OMC.

Contudo, da mesma maneira que Subramanian & Wei (2007), observa-se que os resultados obtidos sugerem que a OMC promoveu o comércio de forma diferente entre países desenvolvidos e em desenvolvimento. Conforme já salientando, a separação desses grupos é necessária, dadas as assimetrias existentes na OMC, que anteriormente foram apresentadas. No entanto, enquanto os autores encontraram que a OMC teve maior êxito em alavancar o comércio entre países desenvolvidos e membros da instituição, entre 1950 e 2000, o resultado estimado aqui diverge dos encontrados por Subramanian & Wei (2007), dado que a *dummy omcd_omcd* não foi significativa. Assim, para o período analisado, de 1990 a 2009, a OMC não teria aumentado o comércio entre os membros desenvolvidos. Esse resultado pode estar ligado ao período analisado nesse artigo, pois é provável que os países desenvolvidos tiveram maiores benefícios com liberalização comercial e, até mesmo, um incremento no comércio bilateral, num período anterior ao analisado, ainda sob a regência do GATT. Faz sentido pensar nisto, pois os países desenvolvidos foram os membros pioneiros do GATT e foram eles que lideraram, durante muitos anos, as negociações multilaterais, liberalizando setores nos quais eles tinham maior interesse. Por sinal, os países desenvolvidos

Tabela 5. Testes de Especificação do Modelo para Painel.

Testes	<i>p</i> -valor
Teste de diferença entre os pares: $F(2.651, 50.355)$	0,000
Teste das <i>dummies</i> de tempo: $F(19, 2.651)$	0,000

Tabela 6. Resultados do modelo.

Variáveis	Coefficiente	σ	Variáveis	Coefficiente	σ
<i>lpib_j</i>	0,8077	0,0530*	<i>apc_nomc</i>	-0,3301	0,0802*
<i>lpib_i</i>	0,6566	0,0600*	<i>omcd_omcd</i>	0,0332	0,0443
<i>lpop_j</i>	-1,1486	0,2273*	<i>omcd_omcs</i>	0,2867	0,0639*
<i>lpop_i</i>	-0,2355	0,3007	<i>omcd_nomc</i>	0,0770	0,0770
<i>Merc</i>	0,9540	0,2514*	<i>omcs_omc</i>	0,2772	0,0715*
<i>merc_imp</i>	0,5683	0,0829*	<i>omcs_nomc</i>	0,0859	0,0837
<i>merc_exp</i>	-0,1470	0,0701**	<i>nomc_nocm</i>	-0,2645	0,1391
<i>nafta</i>	0,3261	0,0814*	<i>da1991</i>	-0,0375	0,0082*
<i>nafta_imp</i>	0,1248	0,0706	<i>da1992</i>	-0,0641	0,0161*
<i>nafta_exp</i>	-0,1484	0,0401*	<i>da1993</i>	-0,0731	0,0231*
<i>can</i>	0,2648	0,3510	<i>da1994</i>	-0,0786	0,0233*
<i>can_imp</i>	0,0129	0,1365	<i>da1995</i>	-0,1210	0,0275*
<i>can_exp</i>	-0,0701	0,0735	<i>da1996</i>	-0,1145	0,0292*
<i>asean</i>	0,3087	0,1433**	<i>da1997</i>	-0,0495	0,0291
<i>asean_imp</i>	-0,0937	0,0562	<i>da1998</i>	-0,0236	0,0307
<i>asean_exp</i>	0,0157	0,0794	<i>da1999</i>	-0,0222	0,0330
<i>ec</i>	0,3298	0,0733*	<i>da2000</i>	0,0670	0,0328**
<i>ec_imp</i>	0,0595	0,0610	<i>da2001</i>	0,0402	0,0330
<i>ec_exp</i>	0,1797	0,0650*	<i>da2002</i>	-0,0370	0,0354
<i>efta</i>	0,2318	0,0881*	<i>da2003</i>	-0,0825	0,0395**
<i>efta_imp</i>	0,0359	0,0664	<i>da2004</i>	-0,0759	0,0445
<i>efta_exp</i>	0,1581	0,0651**	<i>da2005</i>	-0,0714	0,0474
<i>pafta</i>	0,5138	0,2969	<i>da2006</i>	-0,0614	0,0502
<i>pafta_imp</i>	-0,0179	0,0575	<i>da2007</i>	-0,1105	0,0568
<i>pafta_exp</i>	-0,0717	0,0887	<i>da2008</i>	-0,1199	0,0620
<i>apc_omc</i>	-0,0388	0,0445	<i>da2009</i>	-0,2679	0,0610*
Nº de observações					53
Wald χ^2 (52)				20.485,30	[0,000]
Teste de Hausman χ^2 (33)				1.305,04	[0,000]
Teste de Hausman Robusto $F(33, 2.651)$				5,02	[0,000]

Nota: Erros robustos clusterizado no par.

* e ** denotam, respectivamente, nível de confiança de 99% e 95%.



tiveram um cronograma de liberalização mais rigoroso do que os demais.¹⁵

Por outro lado, o coeficiente estimado para os países membros da OMC, considerando países importadores como desenvolvidos e exportadores em desenvolvimento, foi positivo e estatisticamente significativo a 99% (*omcd_omcs*). Esse resultado também foi encontrado por Subramanian & Wei (2007). Porém aqui, esse tipo de comércio norte-sul foi o que a OMC deve maior êxito em alavancar, com crescimento médio de 33% ($e^{0,2867} = 1,33$) para o período analisado. Se levar em conta novamente o período analisado, esse resultado também faz sentido, uma vez que os países em desenvolvimento entraram, de fato, nas negociações da OMC posteriormente aos desenvolvidos e os setores de seu interesse só passaram a ser liberalizados na rodada Uruguai. E, durante a rodada Uruguai, houve uma redução do grau de protecionismo dos países desenvolvidos em relação à produtos agrícolas e têxteis, exportados justamente pelos países em desenvolvimento. De modo complementar, o coeficiente da *dummy* associado a países importadores desenvolvidos que são membros da OMC com países exportadores não pertencentes à organização, foi considerado estatisticamente não diferente de zero (*omcd_nomc*). Assim, os resultados indicam mais uma vez houve discriminação no processo de abertura multilateral.

Os coeficientes estimados das *dummies* dos países em desenvolvimento pertencentes à OMC, de maneira semelhante ao dos países desenvolvidos, indicam que houve discriminação entre membros e não membros. O coeficiente de importação entre membros foi positivo e significativo a 99% (*omcs_omc*). Nesse caso, o crescimento médio para o período foi de 32%. Isso pode ser explicado pela significativa liberalização comercial promovida pelos países em desenvolvimento durante a rodada Uruguai, com as tarifas de importação, em alguns casos, declinando em mais de 50%. Já o coeficiente associado à importação de países em desenvolvimento membros da OMC cujo exportador não é membro da organização não foi significativo (*omcs_nomc*). Por fim, era de se esperar que o comércio entre não membros da instituição não aumentasse e, de fato, isto é o que foi observado nesse artigo, dado que a *dummy* não foi significativa (*nomc_nomc*). Esse resultado também foi encontrado por Subramanian & Wei (2007).

De acordo com os resultados obtidos, é possível verificar que as assimetrias testadas no trabalho de Subramanian & Wei (2007) não se sustentaram ao longo do tempo, tendo se invertido no período mais recente, entre 1990 e 2009. Os autores sugeriam que os países desenvolvidos, entre 1950 e 2000, por estarem mais engajados com a ideia do livre comércio e estarem relativamente mais integrados, tiveram um aumento no comércio bilateral maior do que com os membros em desenvolvimento, pois esses não teriam apresentado o mesmo comprometimento com a liberalização. O que se observou, no período 1990–2009, foi um aumento do comércio entre os países em desenvolvimento membros da OMC, dada a sua recente liberalização multilateral e a elevação das importações dos países desenvolvidos dos países em desenvolvimento, fruto da liberalização de produtos sensíveis da pauta exportadora dos países em desenvolvimento.

5. CONCLUSÕES

Os resultados estimados neste artigo também sustentam a ideia de que existe discriminação entre membros e não membros, e que a OMC foi capaz de gerar um incremento significativo no comércio mundial. Contudo, os coeficientes de importação dos membros desenvolvidos e em desenvolvimento sugerem que, durante 1990 a 2009, a OMC foi mais benéfica para o comércio dos países em desenvolvimento. Esse fato pode ter ocorrido devido a maior abertura comercial entre os países desenvolvidos antes da rodada Uruguai, ou seja, o seu grau de protecionismo nos anos 1990 já era pequeno. No período mais recente, especialmente após a rodada Uruguai, a abertura comercial multilateral mais importante ocorreu justamente nos países em desenvolvimento, que apresentavam um maior grau de protecionismo.

Além disso, a liberalização adicional dos países desenvolvidos beneficiou alguns setores nos quais os países em desenvolvimento são grandes exportadores, como produtos agrícolas e têxteis. Isso ajuda a

¹⁵Esse resultado pode estar associado à periodicidade (anual) das estimações, a amostra dos países e a metodologia econométrica utilizada, que não coincidem com os artigos mencionados.

explicar porque a as importações dos países desenvolvidos só aumentou, de forma significativa, dos países em desenvolvimento e porque houve uma elevação do comércio entre os países em desenvolvimento membros da OMC. Além disso, houve o crescimento no número de países participantes, em sua maioria composta de países em desenvolvimento, que receberam um cronograma de liberalização comercial mais rigoroso que os membros antigos.

Portanto, os resultados sugerem que os países em desenvolvimento, no período 1990–2009, foram os que mais se beneficiaram do aumento do comércio mundial promovido pela OMC, ao contrário do que havia ocorrido entre 1950 e 2000. O artigo mostrou que a OMC teve êxito em atingir seu objetivo de ampliar os fluxos comerciais, embora a abertura comercial novamente tenha ocorrido de forma assimétrica entre países desenvolvidos e em desenvolvimento. A diferença é que, nesse período mais recente, devido a maior liberalização comercial dos países em desenvolvimento da OMC, foi nesses países que ocorreu a maior promoção do comércio provocada pela abertura multilateral, ao contrário do que haviam estimado Subramanian & Wei (2007), para um período anterior.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aitken, N. D. (1973). The effect of the EEC and EFTA on European trade: A temporal cross-section analysis. *American Economic Review*, 63(5), 881–92.
- Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review*, 69, 106–116.
- Anderson, J. E., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the Border Puzzle. *American Economic Review*, 93(1), 170–192. doi: 10.1257/000282803321455214
- Baldwin, R., & Taglioni, D. (2006, September). *Gravity for dummies and dummies for gravity equations* (NBER Working Paper N° 12516). National Bureau of Economic Research. doi: 10.3386/w12516
- Bergstrand, J. (1985). The gravity equation in international trade: Some microeconomic foundations and empirical evidence. *The Review of Economics and Statistics*, 67(3), 474–81.
- Bhagwati, J., Krishna, P. & Panagariya, A. (Eds.). (1999). *Trading blocs: Alternative approaches to analyzing preferential trade agreements*. Cambridge, MA: MIT Press.
- Carrere, C. (2006). Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model. *European Economic Review*, 50(2), 223–247.
- Cheng, I.-H., & Wall, H. J. (2004, July). *Controlling for heterogeneity in gravity models of trade and integration* (Working Paper N° 1999-010E). Federal Reserve Bank of St. Louis. Disponível em: <http://research.stlouisfed.org/wp/1999/1999-010.pdf>
- Deardorff, A. (1998). Determinants of bilateral trade: Does gravity work in a neo-classical world? In J. A. Frankel (Ed.), *Regionalization of the world economy* (pp. 7–31). Chicago: University of Chicago Press.
- Egger, P. (2005). Alternative techniques for estimation of cross-section gravity models. *Review of International Economics*, 13(5), 881–891. doi: 10.1111/j.1467-9396.2005.00542.x
- GATT–General Agreement on Tariffs and Trade (1994). *The results of the Uruguay Round of multilateral trade negotiations / Market access for goods and services: Overview of the results*. Geneva: GATT Secretariat.
- Greene, W. H. (2008). *Econometric analysis* (6ª ed.). New Jersey: Pearson Prentice Hall.
- Helpman, E. (1987). Imperfect competition and international trade: Evidence from fourteen industrial countries. *Journal of the Japanese and International Economies*, 1(1), 62–81.
- Helpman, E., & Krugman, P. (1995). *Market structure and foreign*. Cambridge, MA: MIT Press.
- Helpman, E., Melitz, M. & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics*, 123(2), 441–487. doi: 10.1162/qjec.2008.123.2.441
- Krugman, P. (1991). The move toward free trade zones. *Economic Review, Federal Reserve Bank of Kansas City*, 76, 5–25.
- Linnemann, H. (1966). *An econometric study of international trade flow*. Amsterdam: North-Holland.
- Magee, C. S. P. (2008). New measures of trade creation and trade diversion. *Journal of International Economics*, 75(2), 349–362.



- Matyas, L. (1997). Proper econometric specification of the gravity model. *World Economy*, 20(3), 363–368. doi: 10.1111/1467-9701.00074
- OMC–World Trade Organization (2013). *RTA database*. URL: <http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx>. Acessado em 25 de Fevereiro de 2013.
- Poyhonen, P. (1963). A tentative model for the volume of trade between countries. *Weltwirtschaftliches Archiv*, 90, 93–99.
- Recalde, M., Florensa, L. M. & Iturralde, I. (2008). Gravity equation and trade agreements: A different econometric approach. *Revista de Economía y Estadística*, XLVI(2), 83–104.
- Reis, M. d. (2012). *Os efeitos do novo regionalismo sobre o comércio* (Dissertação de mestrado). PPGE, Universidade do Vale do Rio dos Sinos–UNISINOS, São Leopoldo, RS.
- Rose, A. K. (2004). Do WTO members have more liberal trade policy? *Journal of International Economics*, 63(2), 209–235.
- Santos Silva, J., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4), 641–658.
- Silverstovs, B., & Schumacher, D. (2007). *Estimating gravity equations: To log or not to log?* (Discussion Paper of DIW Berlin N° 739). DIW Berlin, German Institute for Economic Research.
- Soloaga, I., & Winters, A. L. (2001). Regionalism in the nineties: What effect on trade? *The North American Journal of Economics and Financ*, 12(1), 1–29
- Subramanian, A., & Wei, S.-J. (2007). The WTO promotes trade, strongly but unevenly. *Journal of International Economics*, 72(1), 151–175.
- Summers, L. (1991). Regionalism and the world trade system. *Economic Review, Federal Reserve Bank of Kansas City*, 76, 295–302.
- The Economist. (2012, Sep 8). Goodbye Doha, hello Bali. *The Economist*. Acessado em 9 de outubro de 2012: <http://www.economist.com/node/21562196>
- Thorstensen, V. (2001). *OMC–Organização Mundial do Comércio: As regras do comércio internacional e a nova rodada de negociações multilaterais* (2ª ed.). São Paulo: Edições Aduaneiras.
- Tinbergen, J. (1962). *Shaping the world economy: Suggestions for an international economy policy*. New York: Twentieth Century Fund.
- Wei, S.-J. (1996, Apr). *Intra-national versus international trade: How stubborn are nations in global integration?* (NBER Working Paper N° 5531). National Bureau of Economic Research. doi: 10.3386/w5531
- Westerlund, J., & Wilhelmsson, F. (2009). Estimating the gravity model without gravity using panel data. *Applied Economics*, 43(6), 641–649. doi: 10.1080/00036840802599784
- Wonnacott, P., & Lutz, M. (1989). Is there a case for more free trade areas? In J. Schott (Ed.), *Free trade areas and U.S. trade policy* (pp. 59–89). Washington: Institute for International Economics.
- Wooldridge, J. M. (2010). *Introdução à econometria: Uma abordagem moderna* [Tradução da quarta edição norte americana]. São Paulo: Cengage Learning.