

IMPACTOS DO PROGRAMA BOLSA FAMÍLIA NA EDUCAÇÃO DAS CRIANÇAS¹

Ana Paula Kern²

Marcel de Toledo Vieira³

Ricardo da Silva Freguglia⁴

Resumo: Este artigo investiga os impactos do Programa Bolsa Família sobre a educação das crianças de 6 a 17 anos, com enfoque sobre a participação, progressão, repetência e abandono escolar no Brasil e grandes regiões. O Bolsa Família é um programa de transferência condicional de renda. Uma das condicionalidades do Programa Bolsa Família é a matrícula e frequência escolar. A avaliação é realizada utilizando o método de diferenças em diferenças, levando em consideração o desenho amostral complexo. Para esse estudo foi construído um painel de dados longitudinal no nível de indivíduo, a partir dos dados AIBF I e AIBF II. O presente estudo consiste em um avanço na literatura por ser o primeiro a acompanhar o mesmo indivíduo ao longo do tempo e considerar o desenho amostral complexo nas estimações. Os principais resultados mostram que o Bolsa Família não afetou a matrícula escolar das crianças de 2005 para 2009, quando separadas por gênero. No entanto, o Bolsa Família aumentou a probabilidade das crianças se maticularem, especialmente em áreas rurais da região Norte/Centro-Oeste. Para progressão e repetência o Bolsa Família impactou as crianças de 15 a 17 anos de idade residentes na área rural da região Nordeste e na área urbana do Sul/Sudeste, elevando a progressão e diminuindo a probabilidade de repetência. O Bolsa Família também diminui a probabilidade das crianças da área rural do Nordeste evadirem a escola.

Palavras Chave: Bolsa Família, Educação, Condicionalidades, Diferenças em Diferenças, Amostragem Complexa.

Abstract: This paper investigates the impacts of the Bolsa Família Program on the education of children from 6 to 17 years old, focusing on participation, progression, repetition and dropout in Brazil and major regions. The Bolsa Família program is a conditional cash transfer program. One of its conditionalities of the Bolsa Família Program is enrollment and school attendance. The evaluation was performed using the differences in differences method, taking into account the complex sampling design. For this study, a longitudinal data panel at the individual level was constructed from AIBF I and AIBF II data. The present study consists of an advance in the literature because it is the first that accompanies the same individual over time and considers the complex sampling design in the estimations. The main results show that Bolsa Família did not affect the school enrollment of children from 2005 to 2009, when separated by gender. However, Bolsa Família increased the likelihood of children enrolling, especially in rural areas of the North / Central-West region. For progression and repetition, Bolsa Família impacted children aged 15 to 17 residing in the rural area of the Northeast region and in the urban area of the South / Southeast, increasing progression and decreasing the probability of repetition. The Bolsa Família also reduces the probability of children in rural areas of the Northeast to evade school.

Keywords: Bolsa Família, Education, Conditionalities, Differences in Differences, Complex Sampling.

Código JEL: I21, I26, I38

Área 12 - Economia Social e Demografia Econômica

¹ Os autores agradecem o apoio financeiro do CNPq, da Capes e da Fapemig para elaboração desse trabalho.

² Doutoranda em Economia -Universidade Federal de Juiz de Fora, e-mail: ana.kern@yahoo.com.br

³ Professor do Programa de Pós-Graduação em Economia e do Departamento de Estatística - Universidade Federal de Juiz de Fora, email: marcel.vieira@ice.ufjf.br

⁴ Professor da Faculdade de Economia, Coordenador do Programa de Pós-Graduação em Economia e bolsista de produtividade do CNPq - Universidade Federal de Juiz de Fora, email: ricardo.freguglia@ufjf.edu.br

1. INTRODUÇÃO

O objetivo deste estudo é avaliar o impacto do Programa Bolsa Família (PBF) sobre a condicionalidade de matrícula escolar e sobre os indicadores educacionais de progressão, repetência e evasão escolar, de crianças entre 6 e 17 anos de idade, utilizando um painel de dados longitudinal inédito, construído a partir das pesquisas de Avaliação de Impacto do Programa Bolsa família (AIBFI e AIBFII).

A análise da correlação entre a inclusão escolar e o incentivo seletivo proporcionado pela transferência de renda do Programa Bolsa Família ganha relevância na agenda das políticas sociais, apesar da universalização do ensino fundamental no Brasil. O que ocorre é que crianças e jovens que se encontram fora das escolas ou com dificuldades de permanência compõem os estratos populacionais mais vulneráveis em função da sua condição socioeconômica de extrema pobreza, constituindo-se como o público-alvo preferencial dos programas sociais focalizados como o PBF.

Programas de transferência condicional de renda (PTRs) são vistos como novos dispositivos favoritos antipobreza do mundo. Em parte, isto reflete o duplo objetivo de prestar assistência às famílias pobres no curto prazo, e quebrar o ciclo intergeracional da pobreza, criando incentivos para investimento em capital humano nas crianças. Uma atenção particular tem sido dada aos seus impactos sobre a escolaridade, especialmente entre crianças do sexo feminino. PTRs tipicamente incluem uma exigência, ou condicionalidade, de que crianças se matriculem e frequentem a escola e alguns, como o Programa Oportunidades no México, fornecem pagamentos mais elevados para frequência escolar de meninas. O foco de gênero na educação do Programa Oportunidades refletiu em importantes ganhos sociais obtidos através do aumento da escolaridade das meninas que possuíam baixos níveis de resultados escolares. Segundo Levy (2007), no momento inicial do projeto e implementação do Oportunidades, as meninas no México possuíam menores taxas de matrícula e frequência escolar.

Uma extensa literatura documenta que na América Latina, programas PTRs com condicionalidades na educação têm levado a aumentos na taxa de matrícula escolar (Glewwe e Kassouf, 2012; Schaffland, 2012; Aguilar, 2010; Behrman e Parker, 2010; Fiszbein e Schady, 2009). Impactos sobre a matrícula escolar são muitas vezes maiores entre as crianças de famílias pobres (Filmer e Schady, 2008; Maluccio e Flores, 2004) e entre transição de níveis de ensino (De Brauw *et al.*, 2012; Schady *et al.*, 2008). Vários estudos também relatam impactos maiores sobre a escolaridade das meninas do que a para os meninos (De Brauw *et al.*, 2015; Ahmed *et al.*, 2007; Schultz, 2004).

O PBF, criado em 2003, é hoje o maior programa PTR do mundo, com mais de 14 milhões de famílias beneficiárias (MDS, 2017). Inicialmente, a condicionalidade de escolaridade do PBF cobria crianças de 6 a 15 anos, mas um programa complementar introduzido em 2008, chamado Benefício Variável Jovem (BVJ), ampliou o acesso às transferências com condicionalidades às crianças de 16 e 17 anos. Embora o esquema de benefícios pague mais para crianças mais velhas com o BVJ do que para crianças de 6 a 15 anos, essas transferências não fazem distinção de gênero.

Uma série de estudos tem examinado o impacto do PBF nos resultados para escolaridade. Glewwe e Kassouf (2008) utilizam dados do Censo Escolar, de 1998 a 2005, período em que o Bolsa Escola e Bolsa Família foram unificados, para estimar os impactos do programa sobre os resultados do ensino primário. Usando o Censo e aproveitando a variação no momento da introdução do Bolsa Família em todos os locais, eles desenvolveram um modelo do impacto do PBF sobre a matrícula no nível da escola, controlando efeitos fixos para o Estado e para o tempo. Suas estimativas de impacto são identificadas sob a suposição de que, depois de controlar esses efeitos, a presença de BF não está correlacionada com fatores não observados que afetam as decisões de matrícula. Eles encontram que o programa aumenta as matrículas escolares em 5,5-6,5 pontos percentuais, reduz a taxa de abandono em 0,4-0,5 pontos percentuais e aumenta a taxa de aprovação em 0,3-0,9 pontos percentuais. Os resultados são significativos ao nível de significância de 1%.

Simões (2012) utiliza dados do Censo e da Prova Brasil para 2007, com OLS e não encontra nenhum impacto sobre a progressão ou abandono de crianças na área urbana da quarta série. Usando dados da PNAD de 2004 e 2006 e um modelo de escores de propensão com efeito médio do

tratamento sobre os tratados (ATT), Schaffland (2012) também encontra um aumento no número de matrículas de crianças entre 6 e 9 anos e 10 e 7 anos, de aproximadamente 4 pontos percentuais, com efeitos comparáveis encontrados para meninos e meninas.

De Brauw *et al.* (2015), utilizam a pesquisa AIBF I e II, para os anos de 2005 e 2009, para medir o impacto do BF sobre a escolaridade de crianças de 6 a 14 anos e 15 a 17 anos, e desagregam por sexo e região. Eles utilizam o método de escores de propensão ponderado e estimam o efeito de tratamento médio sob os tratados (ATT) e encontram que os efeitos positivos do BF estão concentrados entre as crianças de 15 a 17 anos. Eles também encontram maiores efeitos do programa para zonas rurais do que para zonas urbanas e são maiores para meninas do que para meninos.

O painel de dados utilizado para as análises desse estudo foi construído através da pesquisa de Avaliação de Impacto do Programa Bolsa Família (AIBF), uma base de dados que têm como objetivo acompanhar beneficiários do PBF, desenvolvida sob o comando do Ministério do Desenvolvimento Social, em 2005 e 2009, respectivamente. Essa base de dados é pouco explorada na literatura pela dificuldade de conexão dos dados coletados nos dois períodos. Isso se deve em razão da mudança de instituições responsáveis pela elaboração, aplicação dos questionários e tabulação dos dados, que gerou codificações diferentes para as variáveis nas duas rodadas. Portanto, os estudos que utilizam essa base de dados (AIBF), fazem análises somente para um dos dois anos separadamente, e alguns comparam os resultados dos mesmos.

Sendo assim, o presente trabalho apresenta uma inovação metodológica, construindo um painel de dados longitudinal no nível de indivíduo, utilizando os dados AIBF I e AIBF II, tarefa considerada impossível até o momento (De Brauw *et al.*, 2015). Isso foi possível utilizando-se de técnicas estatísticas avançadas e um esforço rigoroso em identificar o mesmo indivíduo nas duas rodadas, além de identificar as mesmas variáveis, pois elas possuíam nomes e seções diferentes nos dois questionários. Além disso, este trabalho representa mais um avanço na literatura a respeito de programas de transferência de renda, por fazer todas as análises de impacto levando em consideração o desenho amostral complexo, que é outra característica importante normalmente ignorada. Em pesquisas por amostragem, ignorar características do desenho amostral pode enviesar a estimação do impacto e produzir estimativas incorretas, ou seja, pode-se estar encontrando algum impacto onde na verdade não existe e vice-versa (Filho, 2017).

Os dados do painel a nível individual, também permitem estudar o impacto sobre as crianças mais velhas, após a introdução do BVJ, ao contrário da maioria das avaliações que utilizam dados anteriores a 2008. Os dados incluem informações sobre a participação escolar das crianças e níveis de ensino em anos anteriores e atuais, permitindo assim uma análise mais rica tanto na matrícula quanto na transição ao longo do tempo. A amostra grande e de ampla cobertura do programa permite a desagregação por sexo, idade, região e localização. As análises também são segregadas em duas dimensões de heterogeneidade de impacto que são raras em programas PTR na América Latina: (1) comparações entre as três grandes regiões, Norte/Centro-Oeste, Nordeste e Sul/Sudeste, (2) comparações entre áreas urbanas e rurais, e (3) comparações entre crianças dentro da idade obrigatória para o ensino (6-14) e crianças mais velhas (15-17).

O artigo se divide da seguinte forma: na seção 2 encontra-se a descrição do PBF, contextos da política educacional e indicadores educacionais. A seção 3 descreve os dados. A seção 4 apresenta a estratégia empírica. Na seção 5 são discutidos os resultados e finalmente na seção 6, as conclusões.

2. O PROGRAMA BOLSA FAMÍLIA E A EDUCAÇÃO

2.1. O PBF

O Programa Bolsa Família (PBF) foi criado em outubro de 2003, através da unificação de programas já presentes no sistema de proteção social desde a segunda metade da década de 1990, como o Programa Nacional de Renda Mínima vinculada à Educação (Bolsa Escola), o Programa Nacional de Acesso à Alimentação (PNAA), o Programa Nacional de Renda Mínima vinculada à saúde (Bolsa Alimentação) e o Programa Auxílio-Gás. Em 2005, o Programa de Erradicação do Trabalho Infantil

(PETI) também foi incorporado ao PBF. O PBF é um programa de focalização nas famílias consideradas pobres e extremamente pobres.

Transferências associadas ao PBF consistem em (1) pagamento condicional variável por criança com idade entre 0 e 15 anos, por até 5 filhos, para famílias “pobres”, com renda mensal por pessoa entre R\$85,01 e R\$ 170,00; e (2) uma transferência incondicional adicional a famílias “extremamente pobres”, que possuem uma renda mensal por pessoa de até R\$85,00. A transferência está condicionada à mulheres grávidas que fazem pré-natal e cuidados pós-natal, crianças com idades entre 0 e 5 anos estarem em dia com as vacinas e receberem visitas de monitoramento de crescimento, e que todas as crianças entre 6 e 15 anos estejam matriculadas e frequentem a escola pelo menos 85% dos dias de aula.

Em 2008, o programa complementar BVJ foi introduzido, que acrescentou pagamentos variáveis e uma condicionalidade escolar para crianças de idade entre 16 e 17 anos, exigindo que essas crianças participem pelo menos 75% dos dias de escola. Após a introdução do BVJ, todas as famílias participantes do BF com crianças entre 16 e 17 anos foram automaticamente contempladas com o BVJ, sem qualquer procedimento de inscrição adicional.

Os limiões da renda por pessoa e os valores dos benefícios para o BF e BVJ mudaram ao longo do tempo, por conta da inflação. No início do programa, em 2003, os limiões da renda mensal por pessoa eram de R\$100,00 para famílias pobres e R\$50,00 para famílias extremamente pobres. O valor variável condicional pago para crianças de 6 a 15 anos de idade era de R\$15,00 (por criança), e a transferência incondicional para famílias extremamente pobres era de R\$50,00. A tabela 1 mostra os pagamentos do benefício em 2017.

Tabela 1 - Plano de Benefícios para Pagamento do BF e BVJ, em 2017

Renda mensal domiciliar por pessoa	Benefício básico (incondicional)	Benefício variável (pago para crianças de 6 a 15 anos de idade que cumpram as condicionalidades, por até 5 crianças)	Benefício BVJ (pago para crianças de 16 e 17 anos de idade que cumpram as condicionalidades, por até 2 crianças)
Até R\$85,00 (extremamente pobres)	R\$ 85,00	R\$39,00	R\$46,00
Entre R\$85,01 e R\$170,00 (pobres)	-	R\$39,00	R\$46,00

Fonte: Ministério do Desenvolvimento Social

Os pagamentos para crianças de 16 e 17 anos (R\$46,00) são maiores que os pagamentos para crianças de 6 a 15 anos (R\$39,00), mas nenhum tipo de pagamento faz distinção entre o sexo da criança.

As condicionalidades do PBF estão entre os pontos mais polêmicos do programa. Há àqueles que defendem que as condicionalidades são muito importantes ou até mais importantes que a própria transferência de renda, pois são elas que reforçam o direito da criança à educação, ao mesmo tempo que permite a ruptura do ciclo intergeracional da pobreza. Há uma cobrança por parte dos que defendem essas condicionalidades para que o Governo fiscalize com mais rigor o cumprimento dessas contrapartidas e, alguns, até sugerem a criação de outras condicionalidades ou de outros estímulos (por exemplo, um bônus pelo desempenho). Por outro lado, há àqueles que acham que o objetivo principal do BF é o da proteção social, e ao se cobrar excessivamente contrapartidas o objetivo principal se atenua, uma vez que as famílias mais vulneráveis são, provavelmente, as que terão mais dificuldades em cumprir contrapartidas mais rigorosas (Soares e Sátyro, 2009).

Nota-se uma mudança clara por parte do Estado quanto à cobrança do cumprimento das condicionalidades. De acordo com Soares e Sátyro (2009) até setembro de 2006 não havia uma fiscalização efetiva das contrapartidas. A partir de outubro desse mesmo ano, o MDS, o MEC e o Ministério da Saúde implementaram um sistema de acompanhamento do cumprimento das

condicionalidades⁵. Enquanto em novembro de 2006 apenas 62,8% dos beneficiários eram fiscalizados em relação ao cumprimento das contrapartidas, em novembro de 2009 esse percentual já era de 89,6% para as crianças entre 6 e 14 anos e de 79,5% para os jovens entre 15 e 17 anos.

Em caso de descumprimento das condicionalidades, a família recebe advertências e é submetida a diversas avaliações que podem ocasionar a suspensão temporária do benefício e após um ano de descumprimento das contrapartidas, o benefício é cancelado e repassado a outra família. No entanto, o cancelamento do benefício é uma exceção, pois o descumprimento dos compromissos do programa de forma reiterada pode ser um sinal de que a família está em maior vulnerabilidade.

2.2. CONTEXTOS DA POLÍTICA EDUCACIONAL

A tabela 2 apresenta a estrutura da educação para o ensino fundamental (dividido em 2 fases) e o ensino médio. A estrutura da escolaridade mudou em novembro de 2005 para incluir o jardim de infância como obrigatório. Antes da alteração, a escolaridade era obrigatória entre as idades de 7 a 14 anos. A classe de alfabetização era opcional, separada do ensino primário, antes das séries de 1ª a 8ª, e a idade mínima para iniciar o 1º grau era 7 anos.

Tabela 2 - Estrutura da Educação, antes e depois de Novembro de 2005

Antes de Novembro de 2005			Depois de novembro de 2005		
Idade	Curso	Série	Idade	Curso	Série
6	Classede Alfabetização	CA	6	Ensino Fundamental I	1º ano
7-10	Ensino Fundamental I	1ª- 4ª série	7-10		2º-5º ano
11-14	Ensino Fundamental II	5ª- 6ª série	11-14	Ensino Fundamental II	6º-9º ano
15-17	Ensino Médio	9ª-11ª série	15-17	Ensino Médio	10º-12º ano

Fonte: Ministerio da Educação

Após a mudança, a escolaridade se tornou obrigatória a partir dos 6 anos de idade. A classe de alfabetização foi incluída como parte da escola primária e foi renomeada de 1º ano e tornou-se obrigatória antes das renomeadas 2º e 9º ano, e a idade mínima para iniciar o 1º ano passou a ser 6 anos. Se a criança entrou na escola no tempo certo, essas idades obrigatórias corresponderiam à conclusão de ambas as fases do ensino fundamental.

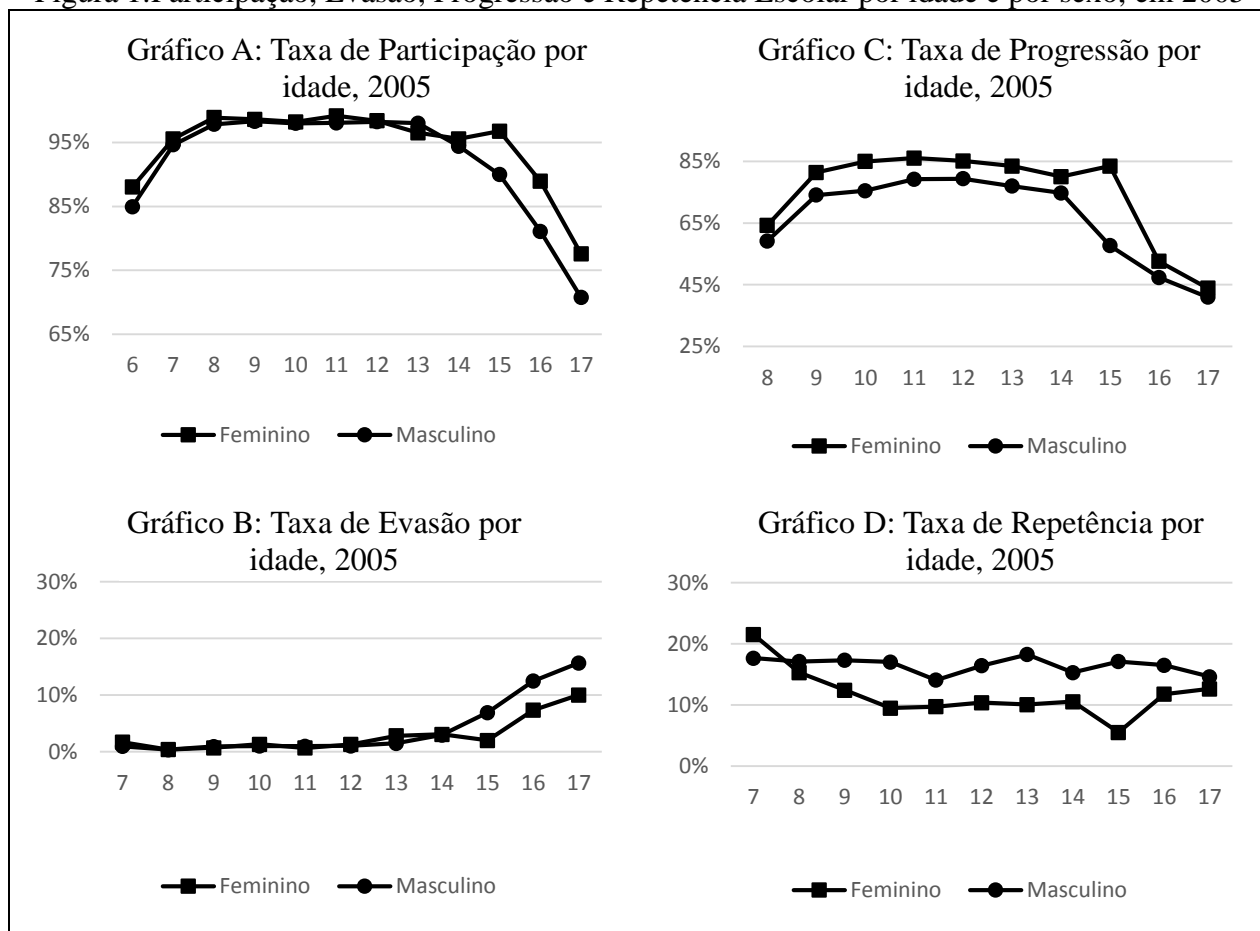
2.3. INDICADORES EDUCACIONAIS

Para interpretar as estimativas de impacto, é útil em primeiro lugar observar os padrões escolares para todas as crianças da amostra da pesquisa em 2005, ou seja, para crianças de 6 a 17 anos. Inicialmente foi considerada a participação na escola incondicional. Na figura 1, o gráfico A mostra que entre meninos e meninas, as taxas de participação na escola tendem a ser bastante elevadas entre as idades de 7 a 14 anos, mas em seguida, observa-se um declínio para as idades de 15 a 17 anos (que corresponde às idades após a escolaridade obrigatória). Os gráficos B, C e D mostram o padrão das taxas de abandono, taxas de progressão e as taxas de repetência escolar das crianças, sob a condição da criança estar matriculada na escola no anterior. O gráfico B mostra que as taxas de abandono se

⁵ Em 2006, o MEC construiu e disponibilizou aos estados e municípios o novo sistema de acompanhamento da frequência escolar do PBF, o Sistema Presença. O novo sistema caracteriza-se pela fácil operacionalização e dispensa o uso de aplicativos, uma vez que foi construído em plataforma web. Entre as inovações apresentadas, destaque para os registros on-line e pela possibilidade que se abriu em descentralizar o registro das informações até a unidade escolar, com a criação do perfil “operador diretor de escola”. O Sistema Presença possui mais de 22 mil usuários cadastrados em todo o país” (CURRALERO *et al.*, 2010).

parecem com o padrão das taxas de participação, com abandono relativamente baixo entre as idades de 7 a 14 anos, e mais elevado para as idades entre 15 e 17 anos. O gráfico C mostra que as taxas de progressão escolar aumentam ligeiramente e em seguida diminuem entre as idades de 7 a 14 anos e também caem para as idades de 15 a 17 anos. O gráfico D mostra que as taxas de repetência são mais ou menos semelhantes após os 9 anos de idade.

Figura 1: Participação, Evasão, Progressão e Repetência Escolar por idade e por sexo, em 2005



Fonte: Elaboração do autor com base nos dados da amostra AIBF I e AIBF II.

Em todas as quatro análises (gráficos A, B, C e D) as meninas tendem a ter melhores resultados do que os meninos da mesma idade. As taxas de participação na escola das meninas são semelhantes ou ligeiramente mais elevadas do que as taxas dos meninos, particularmente em idades mais avançadas. As taxas de desistência das meninas são semelhantes ou ligeiramente menores, particularmente em idades mais avançadas, e as meninas são consistentemente mais propensas a progredirem de série e menos propensas a repetir o ano do que os meninos.

3. DADOS

A AIBF foi criada com o propósito de avaliar o efeito do PBF sobre seus beneficiários ao longo do tempo, comparando-os com os não beneficiários em situação semelhante. Até o presente momento, a pesquisa AIBF teve duas rodadas realizadas, uma em 2005 e outra em 2009, denominadas pesquisa AIBF I (Avaliação de Impacto do Bolsa Família I) e pesquisa AIBF II (Avaliação de Impacto do Bolsa Família II), respectivamente.

A coleta de dados para a avaliação do Programa Bolsa Família seguiu o procedimento no qual a amostra de domicílios foi distribuída em proporções desiguais, segundo três estratos. O primeiro estrato é formado pelos domicílios cujas famílias são beneficiárias do PBF, designados “casos” e, portanto, constituem o grupo de tratamento, denominado “grupo T”. O segundo estrato é constituído pelos

domicílios cujas famílias estão listadas no Cadastro Único, mas não são beneficiárias do PBF (podendo ser beneficiárias de outros programas de transferência de renda), denominados “controles tipo 1” ou “grupo C1”. Por fim o terceiro estrato congrega os domicílios cujas famílias não são cadastradas no CadÚnico e não são beneficiárias, denominados “controles tipo 2” ou “grupo C2” (Sumário Executivo AIBF, 2007).

As diferenças entre ambas as rodadas vão além da data de ocorrência, passando pelo desenho do questionário e até mesmo pela metodologia de gerenciamento dos dados. Isso se deve ao fato de que diferentes instituições foram responsáveis pela execução das duas rodadas da pesquisa. A primeira rodada ficou a cargo do Centro de Desenvolvimento e Planejamento Regional (Cedeplar) da Universidade Federal de Minas Gerais (UFMG) e a segunda pelo Consórcio Instituto Internacional de Pesquisa sobre Políticas Alimentares – IFPRI/Datamétrica Consultoria, Pesquisa e Telemarketing Ltda.

Na rodada de 2005 (AIBF I), o tamanho da amostra foi definido de forma a garantir representatividade para três grandes áreas do País: a região Nordeste (NE); as regiões Sudeste e Sul (SE-SUL) em conjunto; e as regiões Norte e Centro-Oeste (NO-CO), também em conjunto. A coleta domiciliar ocorreu em 269 municípios distribuídos em 24 unidades federativas brasileiras, cobrindo um total de 15.426 domicílios. Com este total, a amostra foi distribuída em 30% de domicílios do grupo T, 60% de domicílios do grupo C1 e 10% de domicílios do grupo C2.

O plano amostral empregado para a pesquisa AIBF foi amostragem dupla (Kish, 1965). Na primeira fase, foi utilizada amostragem conglomerada em uma ou duas etapas para seleção de setores censitários, com estratificação. Na segunda, foi feita seleção de domicílios por amostragem estratificada simples (Cedeplar, 2005).

Na primeira fase, o plano amostral adotado teve estratificação por área geográfica e por tamanho do município. A estratificação dos municípios por tamanho foi feita em dois grupos: os 41 maiores municípios do País, segundo os dados de população do Censo 2000, foram alocados em um estrato de “municípios grandes”. O segundo grupo, composto por todos os demais municípios, foi denominado de “municípios pequenos”.

No estrato dos municípios grandes, o plano amostral foi conglomerado em uma etapa, com seleção de setores censitários como unidades primárias de amostragem (UPAs). A seleção dos setores foi estratificada por município, e realizada por amostragem sistemática com probabilidades proporcionais ao tamanho (método PPT Sistemático; Kish, 1965). A medida de tamanho foi definida como função do total de responsáveis com renda menor que dois salários mínimos em cada setor censitário, conforme os dados do Arquivo Agregado de Setores do Censo 2000. Assim, foi dada maior probabilidade de seleção a setores com maior quantidade de responsáveis pobres, conferindo um efeito adicional de estratificação implícita pelo nível de intensidade da pobreza (Cedeplar, 2005).

No estrato dos municípios pequenos, o plano amostral foi conglomerado em duas etapas. Antes da seleção das unidades conglomeradas, a população foi estratificada segundo as três grandes áreas geográficas citadas. As UPAs foram obtidas mediante a formação de grupos de municípios contíguos, com população mínima de 50.000 habitantes em cada grupo. No conjunto das três áreas, foram formados 1.420 grupos de municípios que serviram como UPAs nesta parte do plano amostral, cuja formação buscou obter a maior heterogeneidade possível em relação a características descritivas dos municípios. As UPAs assim obtidas foram estratificadas, dentro de cada grande área, em três estratos definidos em função da proporção da população atendida por programas de transferência de renda, formando um total de nove estratos de UPAs para fins de amostragem nos municípios pequenos. A alocação da amostra nestes estratos não foi proporcional e buscou alocar maior proporção de UPAs pobres na amostra. A seleção de UPAs foi feita usando amostragem com probabilidades proporcionais ao tamanho através do método Poisson Sequencial, e a medida de tamanho definida de forma semelhante à utilizada para o estrato dos municípios grandes. Uma vez selecionadas as UPAs, os setores foram selecionados no segundo estágio de conglomeração, configurando assim as unidades secundárias de amostragem (USAs), através do método PPT Sistemático. A medida de tamanho empregada foi idêntica à utilizada na seleção de setores no estrato dos municípios grandes (inclusive a

fórmula funcional). Antes de selecionar os setores, estes também foram ordenados segundo a “proporção de responsáveis pobres” em cada setor.

Em 2009, a pesquisa posterior (AIBF II), procurou pesquisar as mesmas famílias. Em AIBF II, 11.433 dos domicílios da amostra da linha de base foram reentrevistados, o que implica uma taxa anual de atrito de aproximadamente 6,5%. As principais fontes de atrito foram devido às equipes de campo serem incapazes de localizar fisicamente endereços e famílias que não residiam mais no endereço registrado (De Brauw *et al.*, 2012). Nessa rodada, os pesos amostrais foram recalculados, levando em consideração o atrito encontrado.

Nas duas rodadas da pesquisa, os dados foram divididos em cinco arquivos com suas respectivas informações, sendo eles: (1) Características dos Domicílios; (2) Alimentos e Bebidas; (3) Características dos Indivíduos; (4) Gastos Coletivos; e (5) Benefícios. A partir dessas duas rodadas da pesquisa, um grande esforço foi feito para a construção de um painel de dados longitudinal a nível individual. Devido a diferenças na codificação das variáveis nas duas rodadas, tal tarefa foi considerada por outros autores como impossível de ser realizada (De Brauw *et al.*, 2015).

Para a construção do painel, em primeiro lugar buscou-se juntar em cada rodada, os cinco arquivos mencionados anteriormente. Para unir os arquivos, fez-se necessário garantir que os assuntos abordados por cada um deles pudessem ser referenciados da mesma maneira, ou seja, que tivessem todas as características por domicílio e cada pessoa em uma única linha. Alguns arquivos, como o caso dos gastos coletivos, alimentos e bebidas e benefícios, apresentavam os itens em suas linhas e nas colunas as caracterizações dos itens agrupados por domicílio.

Além disso, no arquivo de alimentos e bebidas, existiam diferentes números de itens para cada domicílio. O questionário possuía 65 itens, mas os indivíduos podiam responder sobre consumo de outros itens que não constavam no questionário. Portanto, perceberam-se famílias com apenas 65 itens, assim como famílias com até 109 itens.

Para que as linhas desses arquivos se referenciassem a cada domicílio, foi necessária a transposição destas linhas em colunas, criando assim novas variáveis contendo informações para cada item, tendo como referência para transposição o código identificador do domicílio em conjunto com o código identificador do item de interesse. Sendo assim, cada linha passou a ser um domicílio e cada coluna um item. A partir daí os arquivos da base de dados foram unidos formando um arquivo único com todas as informações para 2005. O mesmo foi feito para 2009.

Com dois arquivos somente, contendo todas as observações para 2005 e 2009 respectivamente, foi necessário a padronização das variáveis nos dois bancos, para que a construção do painel fosse possível. As variáveis na rodada de 2009 tinham nomes diferentes da rodada de 2005. No entanto, observou-se que a construção dos nomes das variáveis levava em consideração a seção a que pertencia a pergunta que representavam, a ordem da pergunta e a natureza da resposta, se matricial ou não, em cada questionário. Algumas perguntas e subitens, como por exemplo, na seção referente aos gastos com saúde, não eram as mesmas em 2005 e 2009, o que representou mais uma dificuldade na definição da correspondência entre as variáveis de cada rodada. Sendo assim, foi necessária uma busca nos dois bancos, com o intuito de identificar qual variável de 2009 era a mesma que em 2005.

Todas as variáveis cujas perguntas apresentavam correspondência entre 2005 e 2009 assumiram nome, formato tipo e tamanho de 2009. Foi preciso trocar o nome de 1001 variáveis em 2005, para que ficassem iguais a 2009. Outra necessidade foi trocar o tipo de algumas variáveis que continham informações do tipo data e hora, com a aplicação de máscaras para esses formatos, pois para 2005, por exemplo, a data continha dia, mês e ano, e para 2009 continha apenas mês e ano.

Após essa padronização das variáveis, o próximo passo foi fundamentar a correspondência entre os indivíduos em cada uma das rodadas. Uma forma de averiguar seria confrontar os nomes dos entrevistados. Esta informação, por ser confidencial, não se encontra nos dados brutos disponibilizados pelo Ministério do Desenvolvimento Social (MDS). Embora tenha sido solicitada aos responsáveis pelo armazenamento da base de dados identificada, esta informação não foi disponibilizada.

Optou-se então por comparar os indivíduos de cada uma das bases confrontando data de nascimento, sexo e idade, dentro do domicílio. Esta estratégia restringiu o número de indivíduos de

68.395 na primeira rodada, para 43.412 indivíduos que foram localizados na base de dados da segunda rodada, o que corresponde a 63,47% da amostra. Para uma análise de robustez, foram realizados testes verificando a frequência relativa e absoluta das variáveis para verificar o efeito desse atrito, chegando à conclusão que essa perda não tornou a amostra enviesada⁶. Algumas variáveis também precisaram ser excluídas, tais como variáveis que só continham informações para um dos dois anos, ou que não possuíam referência em um dos anos. Dessa forma, foi construído um painel de dados balanceado, contendo todos os dados dos indivíduos visitados em ambas as rodadas. Este painel foi utilizado para as análises desse estudo.

A partir do painel construído e informações dos grupos T, C1 e C2, foram construídos o grupo de tratamento e controle para esse estudo. Com o intuito de capturar possíveis mudanças nos indicadores educacionais das crianças de 6 a 17 anos de idade, como grupo de tratamento considerou-se indivíduos que não receberam Bolsa Família em 2005, mas receberam em 2009 e como grupo de controle os indivíduos que não receberam Bolsa Família nem em 2005 e nem em 2009, mas estavam cadastrados no CadÚnico. Assim seria possível capturar o efeito do PBF. Foram considerados os pesos amostrais calculados para amostra de 2009, pois dessa maneira, além de fazer inferência estatística para toda a população, o atrito foi corrigido.

O painel possui informações sobre escolaridade para cada indivíduo, incluindo se o indivíduo está atualmente na escola e o curso/nível correspondente, situação de matrícula do ano anterior e curso/nível, e o maior curso/nível que o indivíduo concluiu se não está matriculado na escola.

Embora a estrutura escolar tenha se alterado de 2005 para 2009, os entrevistados foram instruídos a codificar a rodada de 2009 pela estrutura de 2005, para que houvesse consistência.

Esse artigo se concentra no seguinte conjunto de indicadores com base nas informações para crianças de 6 a 17 anos listadas a seguir: (1) uma medida incondicional de se a criança participar atualmente da escola (“matrícula escolar”); (2) condicional a criança participar da escola no ano anterior, se a criança tem progredido a partir do nível da série anterior (“progressão escolar”); (3) condicional a criança participar da escola no ano anterior, se a criança tem repetido o nível de grau anterior (“repetência”); (4) condicional a criança participar da escola no ano anterior, se a criança saiu da escola (“evasão”).

Como se consegue acompanhar o mesmo indivíduo de 2005 em 2009 consegue-se construir medidas relacionadas com trajetórias escolares. Sendo assim, o trabalho se concentra nos impactos sobre participação (matrícula) escolar e a trajetória do aluno (progressão, repetência e evasão).

4. ESTRATÉGIA EMPÍRICA

Com o objetivo de estimar o efeito do PBF nos resultados escolares das crianças de 6 a 17 anos, utilizou-se o método de Diferenças em Diferenças (DID). Uma das principais vantagens do método DID é que ele é capaz de controlar as influências, sobre a variável de resultado, das características não observáveis dos indivíduos que sejam fixas no tempo (Wooldridge, 2007). Essa é uma vantagem importante do método, pois, muitas vezes, essas características fixas não observáveis influenciam a participação no programa. Assim, o DID é capaz de levar em conta a associação entre a variável de resultado, a participação no programa e as características não observáveis dos indivíduos que sejam invariantes no tempo. Com isso, o DID oferece uma ferramenta útil para contornar um possível problema de viés de autosseleção que possa existir.

Quando os indivíduos em um grupo de tratamento e um grupo de controle são observados nos períodos de pré-tratamento e pós-tratamento e as tendências do tempo de pré-tratamento na variável de resultado não são significativamente diferentes nos dois grupos, o modelo DID pode ser usado para estimar o efeito do tratamento nos tratados. Uma maneira de especificar o modelo é definindo uma variável *Post* que é igual a um se a observação for do período pós-tratamento e zero se for do período de pré-tratamento; e uma variável *Treat* que é igual a um se a observação for do grupo de tratamento e

⁶ Os resultados dos testes encontram-se disponíveis mediante solicitação aos autores.

zero se for do grupo de controle. Deixando o vetor X representar algumas variáveis explicativas adicionais, incluindo um termo constante, o modelo linear DID aparece da seguinte maneira:

$$y = X\beta + \beta_1 Post + \beta_2 Trat + \beta_{12} (Post \times Trat) + u \quad (1)$$

$$E(y|x, Trat=1, Post=1) = (X\beta + \beta_1 + \beta_2 + \beta_{12})$$

$$E(y|x, Trat=1, Post=0) = (X\beta + \beta_2)$$

$$E(y|x, Trat=0, Post=1) = (X\beta + \beta_1)$$

$$E(y|x, Trat=0, Post=0) = (X\beta)$$

Sendo assim, a mudança no valor da variável dependente (y) do período de pré-tratamento para o período de pós-tratamento para o grupo de tratamento é:

$$\left(\frac{\Delta E(X\beta + \beta_1 Post + \beta_2 Treat + \beta_{12}(Post \times Treat))}{\Delta Post} \Big|_{Treat=1} \right) = E(X\beta + \beta_1 + \beta_2 + \beta_{12}) - E(\beta_2)$$

E para o grupo de controle é:

$$\left(\frac{\Delta E(X\beta + \beta_1 Post + \beta_2 Treat + \beta_{12}(Post \times Treat))}{\Delta Post} \Big|_{Treat=0} \right) = E(X\beta + \beta_1) - E(X\beta)$$

Portanto, o efeito DID é:

$$DID = [E(X\beta + \beta_1 + \beta_2 + \beta_{12}) - E(X\beta + \beta_2)] - [E(X\beta + \beta_1) - E(X\beta)] \quad (2)$$

A equação 2 é o resultado do DID linear, que é o mesmo que o efeito do tratamento nos tratados. No modelo linear, o efeito do tempo é constante em ambos os grupos (tratamento e controle). Se o valor da variável dependente (y) sem o tratamento não é limitado, então ele pode ser determinado pelo efeito da interação β_{12} .

No entanto, como afirma Puhani (2012), isso não é verdade para um modelo não-linear, como é o caso das variáveis desse estudo, que são variáveis binárias. Em contraste com um modelo linear (equação 1), o efeito marginal de uma variável explicativa em um modelo não-linear não é constante em todo o seu alcance, mesmo na ausência do termo de interação (isto é, $\beta_{12} = 0$).

A razão é que no modelo linear, a equação 2 simplifica β_{12} , mas isso não é verdadeiro em modelos não lineares. Dito de outra maneira, se o termo da interação $Post \times Treat$ for igual a zero em um modelo linear, então $\frac{\Delta E}{\Delta Post}$ é o mesmo para $Treat=1$ e $Treat=0$. Isso não é verdade para o modelo não linear, porque o movimento de $Treat=1$ para $Treat=0$, por si só, induz uma mudança em $\frac{\Delta E}{\Delta Post}$.

Para isolar a diferença em $\frac{\Delta E}{\Delta Post}$ no grupo de tratamento em um modelo não linear, é necessário calcular o DID mantendo $Treat=1$ ao mudar $Post \times Treat$ de zero para um. Para ver isso, considere o modelo não linear:

$$P(y=1|x) = F(X\beta + \beta_1 Post + \beta_2 Trat + \beta_{12} (Post \times Trat)) + u \quad (3)$$

A mesma lógica do DID pode ser aplicada:

$$P(y=1|x, Trat=1, Post=1) = F(X\beta + \beta_1 + \beta_2 + \beta_{12})$$

$$P(y=1|x, Trat=1, Post=0) = F(X\beta + \beta_2)$$

$$P(y=1|x, Trat=0, Post=1) = F(X\beta + \beta_1)$$

$$P(y=1|x, Trat=0, Post=0) = F(X\beta)$$

O parâmetro β_1 permite que o seu índice linear (e, portanto, o $P(y=1|x)$) seja diferente para todos os indivíduos no período pós-tratamento em comparação com o período de pré-tratamento. β_2 permite que seu índice linear (e, portanto, o $P(y=1|x)$) também seja diferente para os indivíduos tratados em comparação com os indivíduos do grupo de controle. β_{12} permite que o índice linear seja diferente no período pós-tratamento e, portanto, a probabilidade condicional de que $P(y=1|x)$ seja diferente, além da diferença atribuível à não-linearidade do modelo para indivíduos no grupo de tratamento em comparação com o grupo de controle. É essa diferença adicional nas diferenças que fornece uma medida do efeito do tratamento nos tratados para o modelo não linear.

Sendo assim, o resultado DID para o modelo não linear é dado por:

$$\begin{aligned}
 (DID | Treat = 1) &= [F(\beta_1 + \beta_2 + \beta_{12}) - F(\beta_2)] - [F(\beta_1) - F(0)] \\
 &\quad - [F(\beta_1 + \beta_2) - F(\beta_2)] - [F(\beta_1) - F(0)] \\
 &= F(\beta_1 + \beta_2 + \beta_{12}) - F(\beta_1 + \beta_2)
 \end{aligned}
 \tag{4}$$

(Post x Trat=1)
(Post x Trat=0)

Esta fórmula é contra intuitiva porque o terceiro termo entre colchetes implica que $Post=1$ e $Treat=1$, mas $Post \times Treat = 0$. A equação 4 é igual a zero, se e somente se, β_{12} for igual a zero.

A extensão de que a adição de um termo de interação explícita altera os efeitos marginais ou transversais podem ser respondidas pela comparação dos efeitos marginais dos modelos com e sem um termo de interação explícita.

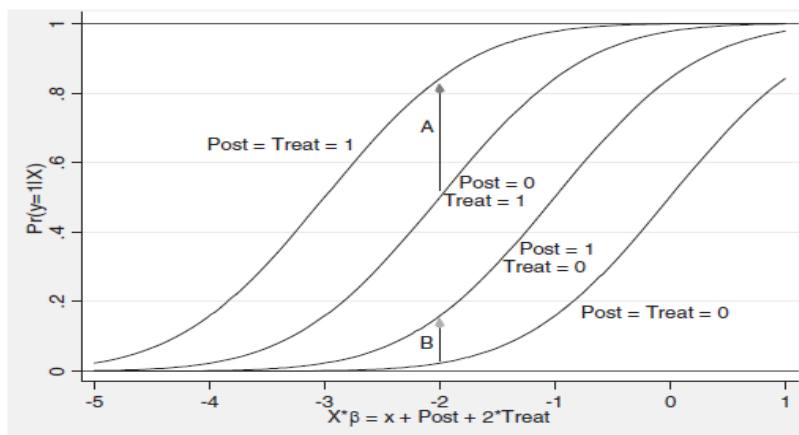
Para ilustrar essa diferença, considere as Figuras 2 e 3, que mostram o efeito de adicionar o termo de interação explícita a um modelo *probit*. A Figura 1 mostra as relações entre a probabilidade condicional de que $y = 1$ e uma variável explicativa X , em um modelo sem o termo de interação:

$$P(y=1|x) = F(X\beta + \beta_1 Post + \beta_2 Treat +) + u$$

Os parâmetros β , β_1 e β_{12} foram ajustados iguais a 1, e β_2 foi ajustado igual a 2 (β_2 precisa ser diferente de β_1 para que as linhas não se encontrem uma sobre a outra). Na Figura 1, a curva mais para a direita corresponde a $Post = 0, Treat = 0$, portanto, a função de distribuição cumulativa da variável X tem uma média igual a zero. A curva à esquerda corresponde a $Post = 1, Treat = 0$. A terceira linha da direita corresponde a $Post = 0, Treat = 1$. Finalmente, a curva mais à esquerda corresponde a $Post = 1, Treat = 1$.

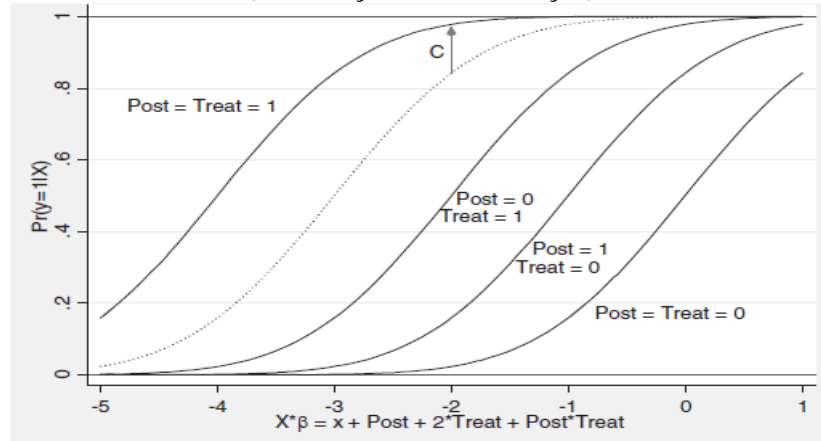
A estimativa DID é a diferença no eixo vertical entre a terceira e a quarta linha da direita (distância A) em relação à primeira e segunda linhas da direita (distância B) avaliada em um valor específico das variáveis explicativas X .

Figura 2 - Modelo não linear (Probit) sem o termo de interação



Extraído de: Mandic, Norton, Dowd, 2012

Figura 3 - Modelo não linear (Probit) com o termo de interação (Diferenças em Diferenças)



Extraído de: Mandic, Norton, Dowd, 2012

A Figura 2 mostra que na ausência de um termo de interação explícita, a distância A e a distância B não são iguais, como seria em um modelo linear e, portanto, o modelo não-linear produzirá uma estimativa DID diferente de zero mesmo sem um termo de interação. Ou seja, a parcela do efeito DID observada em Puhani (2012) deve ser mantida constante ao avaliar a versão DID das interações em modelos não-lineares.

Conforme mostrado na Figura 3, incluindo o termo de interação, a curva mais distante da direita, correspondente a $Post = 1, Treat = 1$, se desloca ainda mais para a esquerda. Esse deslocamento adicional para cima no $P(y=1|x)$ da linha pontilhada para a linha contínua acima (distância C na Figura 3) é a parcela do efeito DID atribuível à interação explícita $Post \times Treat$. Neste caso especial que mantém os dois, $Treat$ e $Post$ constantes, mas permite que $Post \times Treat$ varie, Puhani (2012) mostrou que o efeito do tratamento nos tratados para o modelo não linear não é uma diferença cruzada simples, mas sim a diferença entre as diferenças cruzadas.

No entanto, ao trabalhar com dados amostrais complexos e aplicar diretamente o método de Diferença em Diferenças, ignorando as suas características, pode-se obter estimativas incorretas para as variâncias das estimativas pontuais dos parâmetros de interesse, e conseqüentemente de suas medidas de precisão (testes de significância e intervalos de confiança, por exemplo), o que pode comprometer a qualidade da inferência estatística. A desconsideração do plano amostral pode ainda resultar em estimativas viciadas para os parâmetros pontuais (médias, proporções e coeficientes de modelos de regressão, por exemplo).

Este viés se deve aos efeitos de estratificação, conglomeração e pesos desiguais. Logo, os intervalos de confiança obtidos serão mais largos ou estreitos do que deveriam e os testes de hipóteses realizados serão, respectivamente, mais conservadores ou liberais⁷, incluindo menos ou mais variáveis no modelo a ser ajustado do que o devido (Filho, 2017).

Para medir o efeito do plano amostral na estimação, Skinner, Holt e Smith (1989) propuseram uma medida denominada Efeito do Plano Amostral (EPA) ampliado. O EPA é dado pela variância de um estimador $\hat{\theta}$ calculada sobre a distribuição do plano amostral considerado (também chamada de variância verdadeira) dividida pela estimativa v_0 da variância do estimador $\hat{\theta}$ (Pessoa e Silva, p. 48):

$$EPA(\hat{\theta}) = \frac{V_{Verd}(\hat{\theta})}{E_{Verd}(v_0)}$$

⁷ Na modelagem estatística, um teste de hipóteses é dito liberal se, tomadas k amostras de tamanho igual da mesma população, a taxa de rejeição da hipótese nula (o coeficiente associado à variável sob teste é nulo) pelo teste, realizado em cada uma das k amostras, é maior do que o nível de significância (complementar do nível de confiança) do teste. Um teste é conservador quando ocorre o fenômeno inverso (Filho, 2017).

Desta forma, o $EPA(\hat{\theta}, v_0)$ mede a tendência de v_0 a subestimar ou superestimar $V_{Verd}(\hat{\theta})$, variância verdadeira de $\hat{\theta}$. Quanto mais afastado de 1 for o valor de EPA ($\hat{\theta}, v_0$), maior é a consequência de se ignorar o plano amostral complexo e a especificação correta da estrutura populacional.

5. RESULTADOS

Nesta seção são apresentadas as estatísticas descritivas, assim como o impacto do PBF nos indicadores educacionais para crianças entre 6 e 17 anos de idade.

5.1. ANÁLISES DESCRITIVAS

A Tabela 3 apresenta a proporção de crianças que estão matriculadas na escola e a proporção de crianças que progridem, evadem ou repetem o ano na escola. Além disso, foi calculado o EPA para medir o efeito do plano amostral. Os valores são maiores que 1 e indicam que ao ignorar o plano amostral complexo, pode-se estar encontrando proporções incorretas.

Tabela 3 - Análises Descritivas das variáveis de interesse, com e sem o plano amostral

Variável	Sem o plano amostral				Com o plano amostral				EPA	
	2005		2009		2005		2009		2009	
	T	C	T	C	T	C	T	C	T	C
Matrícula Escolar	0.7108 (0.0058) n= 5987	0.7753 (0.0056) n= 5395	0.8677 (0.0052) n= 4105	0.8479 (0.0062) n= 3256	0.6107 (0.0899) n= 5987	0.7359 (0.0215) n= 5395	0.8627 (0.0159) n= 4105	0.8402 (0.0205) n= 3256	9,35	10,93
Progressão Escolar	0.7411 (0.0074) n= 3434	0.7732 (0.0071) n= 3409	0.8128 (0.0083) n=2159	0.8591 (0.0085) n= 1669	0.7056 (0.0268) n= 3434	0.7671 (0.0151) n= 3409	0.8311 (0.0202) n=2159	0.8487 (0.0192) n= 1669	5,92	5,10
Evasão Escolar	0.0295 (0.0021) n= 5982	0.0296 (0.0023) n= 5394	0.0791 (0.0042) n= 4092	0.0822 (0.0048) n= 3247	0.0222 (0.0055) n= 5982	0.0313 (0.0050) n= 5394	0.0742 (0.0108) n= 4092	0.0784 (0.0008) n= 3247	6,61	0,03
Repetência Escolar	0.2588 (0.0074) n= 3434	0.2267 (0.0071) n= 3409	0.1871 (0.0083) n= 2159	0.1408 (0.0085) n= 1669	0.2943 (0.0268) n= 3434	0.2328 (0.0151) n= 3409	0.1688 (0.0202) n= 2159	0.1512 (0.0192) n= 1669	5,92	5,10

Fonte: Painel produzido a partir da AIBF I e AIBF II. Os valores entre parênteses são os erros padrão.
Nota: T representa o grupo de tratamento e C o grupo de controle. EPA é o efeito do plano amostral.

Pode-se observar que cerca de 61% das crianças beneficiárias do PBF no Brasil estavam matriculadas na escola em 2005 e em 2009 essa porcentagem aumenta para cerca de 84%. Para as crianças não beneficiárias do PBF, essas proporções são de aproximadamente 73% para 2005 e 84% para 2009.

É interessante notar que todos os indicadores melhoram de 2005 para 2009, com exceção da evasão escolar, que para o grupo de crianças beneficiárias foi 2% em 2005, passando para 7% em 2009. O mesmo aconteceu com o grupo de crianças não beneficiárias, que passou de 3% em 2005 para 7% em 2009. No entanto, as proporções de progressão escolar são altas, em média 73% em 2005 e 83% em 2009, enquanto as taxas de repetência escolar são baixas, em média 26% em 2005 e 15% em 2009.

5.2. RESULTADOS DAS ESTIMAÇÕES

Nesse ensaio foram estimados os efeitos do PBF para os indicadores educacionais. As análises também levam em consideração o plano amostral, pois, ao ignorá-lo, podem-se obter conclusões enviesadas, isto é, pode-se encontrar algum impacto quando na verdade o mesmo não existe e vice-

versa. Considerar o plano amostral também faz com que os resultados possam ser representativos para toda a população, ou seja, para todo o Brasil.

Os efeitos da expansão do PBF podem ser heterogêneos, isto é, podem depender das características dos beneficiários. Examina-se essa possibilidade dividindo a amostra de crianças por gênero e por faixa etária. Isso se justifica para que se possa capturar o efeito do PBF tanto nas crianças mais novas, que é a condicionalidade para receber BF, tanto as crianças mais velhas, após a introdução do BVJ. Na tabela 3 encontram-se os resultados, comparando-se áreas rurais e urbanas para crianças do sexo masculino e feminino.

Tabela 4 - Impacto do Bolsa Família nos indicadores escolares, desagregado por idade, sexo e localização com o plano amostral

	Rural			Urbano		
	Todos	Desagregado por sexo		Todos	Desagregado por sexo	
		Meninas	Meninos		Meninas	Meninos
Participação na Escola						
Idades 6-17 anos	0,0552 (0,0404) n= 3178	0,0585 (0,0359) n= 1517	0,0469 (0,0734) n= 1658	0,0210 (0,0170) n= 14827	0,0001 (0,0254) n= 7210	0,0309 (0,0213) n= 7612
Idades 6-14 anos	0,0381 (0,0648) n= 1995	0,1750 (0,2716) n= 940	-0,0048 (0,0454) n= 1053	-0,0047 (0,0041) n= 9703	-0,0074 (0,0055) n= 4777	-0,0018 (0,0055) n= 4923
Idades 15-17 anos	0,0371 (0,0618) n= 669	0,0227 (0,0449) n= 338	0,0463 (0,1396) n= 330	0,0277 (0,0253) n= 3282	-0,0031 (0,0116) n= 1572	0,0839 (0,0727) n= 1710
Progressão na Escola						
Idades 6-17 anos	0,0089 (0,0711) n= 1792	-0,0098 (0,0316) n= 877	0,0567 (0,1279) n= 912	0,0249 (0,0948) n= 8602	0,0199 (0,0339) n= 4262	-0,0283 (0,0994) n= 4337
Idades 6-14 anos	-0,0050 (0,0715) n= 1476	-0,0118 (0,0320) n= 709	0,0200 (0,1228) n= 765	0,0057 (0,0393) n= 6982	0,0341 (0,0452) n= 3498	-0,0316 (0,0575) n= 3481
Idades 15-17 anos	0,2340 (0,2543) n= 417	0,0961 (0,1854) n= 214	0,3778 (0,3695) n= 202	0,0972 (0,1413) n= 1896	-0,0366 (0,1034) n= 909	0,3114 (0,2286) n= 987
Repetência Escolar						
Idades 6-17 anos	-0,0089 (0,0711) N= 1792	0,0098 (0,0316) n= 877	-0,0567 (0,1279) n= 912	-0,0249 (0,0948) n= 8602	-0,0199 (0,0339) n= 4262	0,0283 (0,0994) n= 4337
Idades 6-14 anos	0,0050 (0,0715) n= 1476	0,0118 (0,0320) n= 709	-0,0200 (0,1228) n= 765	-0,0057 (0,0393) n= 6982	-0,0341 (0,0452) n= 3498	0,0316 (0,0575) n= 3481
Idades 15-17 anos	-0,2340 (0,2543) n= 417	-0,0961 (0,1854) n= 214	-0,3778 (0,3695) n= 202	-0,0972 (0,1413) n= 1896	0,0366 (0,1034) n= 909	-0,3114 (0,2286) n= 987
Evasão Escolar						
Idades 6-17 anos	0,0009 (0,0105) N= 3174	-0,0032 (0,0092) n= 1515	0,0103 (0,0318) n= 1656	0,0014 (0,0048) N= 14804	0,0093 (0,0068) n= 7200	-0,0111 (0,0141) n= 7599
Idades 6-14 anos	-0,1221 (0,1669) N= 1994	-0,0967 (0,1897) n= 940	-0,1192 (0,2352) n= 1052	0,0031 (0,0027) N= 9686	0,0051 (0,0044) n= 4771	0,0016 (0,0025) n= 4912
Idades 15-17 anos	-0,0001 (0,0195) N= 669	0,0012 (0,0027) n= 338	-0,0046 (0,0685) n= 330	-0,0039 (0,0057) N= 3274	0,0019 (0,0035) n= 1569	-0,0137 (0,0184) n= 1705

Fonte: Painel produzido a partir da AIBF I e AIBF II. Os valores entre parênteses são os erros padrão.

Nota: * Significativo ao nível de 10%

** Significativo ao nível de 5%

*** Significativo ao nível de 1%

Os resultados mostram que quando se desagrega por localização e gênero, o PBF não tem impacto em nenhum indicador educacional. Esse resultado levanta algumas questões que já foram tratadas pelos pesquisadores na literatura, de que as condicionalidades exigidas por programas de transferência de renda não deviam existir, dada a dificuldade das famílias em cumpri-las, e, de que, a fiscalização do cumprimento dessas condicionalidades não possui caráter punitivo, pois sugere que famílias que não estão cumprindo as condições são as mais vulneráveis.

Assim, autores como Schwartzman (2009), defendem que programas como o Bolsa Família não deveriam possuir condicionalidades, deveriam apenas garantir a transferência de renda àqueles que se encontram em situação de pobreza, uma vez que estas condicionalidades não estão conjugadas com uma política educacional de melhoria da qualidade de ensino, e assim, não quebrariam o ciclo intergeracional da pobreza. Segundo esse autor, as pesquisas mostram que a população valoriza muito a educação, e de fato a entrada e a permanência das pessoas nas escolas vem aumentando ano a ano, independentemente da existência ou não de bolsa família ou subsídio semelhante. O autor diz ainda que, havendo boas escolas e serviços de saúde acessíveis, a população naturalmente buscará estes serviços.

Segundo Sjoberg e Nilsson (2013), o PBF não aumenta a matrícula escolar, uma vez que as crianças teriam sido matriculadas mesmo sem a concessão e as crianças de famílias que foram bloqueadas ainda frequentam a escola.

Quanto a progressão, repetência e evasão escolar, alguns autores como Oliveira e Duarte (2005) e Reimers *et al.* (2006) questionam o fato do PBF (assim como quase todos os PTRs na América Latina) assumir que a participação dos alunos na escola é uma condição suficiente para a aprendizagem, sem considerar o que realmente acontece no interior da escola, principalmente ao se observar a baixa proficiência dos estudantes das escolas públicas brasileiras e que, nesse sentido, o PBF por si só não afeta a progressão escolar de seus beneficiários.

Após apresentar evidências de que o PBF não tem impacto sobre os indicadores educacionais, também foi investigada a possibilidade de impactos examinando se o efeito varia de acordo com a idade das crianças beneficiárias em diferentes regiões geográficas do Brasil. Isso pode ser importante porque as regiões brasileiras são bastante heterogêneas em muitos aspectos de desenvolvimento social e cultural. De acordo com os dados do MDS, a distribuição espacial das transferências do PBF é altamente desigual entre as regiões do país. De fato, o principal destino dos recursos do programa é a região Nordeste (53,2 %), seguida pela região Sudeste (23,4 %). Longe de representar uma falha na distribuição de recursos, isso é resultado do objetivo do programa de reduzir os níveis de pobreza no país, pois, de acordo com o MDS, quase três quartos das famílias pobres no Brasil em 2005 estavam concentrados nessas duas regiões. O impacto (em termos de efeitos marginais) do PBF sobre os indicadores escolares das crianças por região é mostrado na Tabela 5.

Tabela 5 - Impacto do Programa Bolsa Família nos indicadores de Educação das crianças de 6 a 17 anos, desagregado por região e localização, com o plano amostral

Variável	Brasil			Norte/Centro-Oeste			Nordeste			Sul/Sudeste		
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural
Matrícula Escolar												
Idade 6-17 anos	0.0277* (0.0157) n=18487	0.0210 (0.0170) n= 14827	0.0552 (0.0404) n= 3178	0.0662*** (0.0279) n= 5768	0.0539** (0.0233) n= 5314	0.7084*** (0.2159) n= 328	0.0741** (0.0322) n= 6108	0.0774** (0.0400) n=4041	0.0793 (0.0627) n= 1900	-0.0213 (0.0299) n= 6611	-0.0267 (0.0234) n= 5472	-0.0246 (0.4661) n= 950
Idade 6-14 anos	-0.0099 (0.0062) n= 11991	-0.0047 (0.0041) n= 9703	0.0381 (0.0648) n=1995	0.0335 (0.0474) n= 3801	-0.0020 (0.0093) n= 3499	0.9143*** (0.1917) n= 223	0.0011 (0.0072) n= 3836	-0.0010 (0.0065) n= 2560	0.0287 (0.0615) n=1180	-0.0107 (0.0088) n= 4354	-0.0090 (0.0887) n= 3644	-0.0205 (0.0383) n= 592
Idade 15-17 anos	0.0264 (0.0223) n= 4091	0.0277 (0.0253) n= 3282	0.0371 (0.0618) n= 669	0.0984 (0.0777) n= 5314	0.1412 (0.0992) n= 1172	0.0203 (0.1663) n= 74	0.0945 (0.0684) n= 1321	0.1061 (0.0894) n= 869	0.0840 (0.1216) n= 404	0.0118 (0.232) n= 1488	0.0096 (0.0220) n= 1241	0.0363 (0.0908) n= 191
Progressão Escolar												
Idade 6-17 anos	0.0241 (0.0413) n=10671	0.0249 (0.0948) n= 8602	0.0089 (0.0711) n= 1792	-0.1093 (0.0685) n= 3181	-0.1114 (0.0753) n= 2930	-0.0678 (0.3013) n= 179	0.0005 (0.0474) n= 3609	-0.0147 (0.0510) n=2414	0.0827 (0.1678) n= 1098	0.0688 (0.0806) n= 3881	0.0892 (0.0956) n= 3258	-0.0440 (0.0640) n= 515
Idade 6-14 anos	0.0048 (0.0338) n=8665	0.0057 (0.0393) n= 6982	-0.0050 (0.0715) n=1476	-0.0938 (0.0776) n= 2585	-0.0973 (0.0856) n= 2379	-0.0617 (0.3212) n= 153	-0.0208 (0.0381) n= 2877	-0.0356 (0.0428) n= 1925	0.0479 (0.1379) n=881	0.0277 (0.0516) n= 3203	0.0440 (0.0612) n= 2678	-0.0535 (0.0584) n= 442
Idade 15-17 anos	0.1309 (0.1220) n= 2382	0.0972 (0.1413) n= 1896	0.2340 (0.2543) n= 417	0.2281* (0.1292) n= 712	-0.1902 (0.1274) n= 829	0.2655 (0.5829) n= 39	0.1062 (0.2151) n= 823	-0.0923 (0.1475) n= 524	0.7714*** (0.3214) n= 272	0.3666* (0.2117) n= 847	0.5128** (0.2346) n= 717	0.0251 (0.0922) n= 106
Repetência Escolar												
Idade 6-17 anos	-0.0241 (0.0413) n=10671	0.0249 (0.0498) n= 8602	-0.0089 (0.0711) n= 1792	0.1093 (0.0685) n= 3181	0.1114 (0.0753) n= 2930	0.0678 (0.3013) n= 179	-0.0005 (0.0474) n= 3609	0.0147 (0.0510) n=2414	-0.0827 (0.1678) n= 1098	-0.0688 (0.0806) n= 3881	-0.0892 (0.0956) n= 3258	0.0440 (0.0640) n= 515
Idade 6-14 anos	-0.0048 (0.0338) n=8665	-0.0057 (0.0393) n= 6982	0.0050 (0.0715) n=1476	0.0938 (0.0776) n= 2585	0.0973 (0.0856) n= 2379	0.0617 (0.3212) n= 153	0.0208 (0.0381) n= 2877	0.0356 (0.0428) n= 1925	-0.0479 (0.1379) n=881	-0.0277 (0.0516) n= 3203	-0.0440 (0.0612) n= 2678	0.0535 (0.0584) n= 442
Idade 15-17 anos	-0.1309 (0.1220) n= 2382	-0.0972 (0.1413) n= 1896	-0.2340 (0.2543) n= 417	-0.2281* (0.1292) n= 712	0.1902 (0.1274) n= 829	-0.2655 (0.5829) n= 39	-0.1062 (0.2151) n= 823	0.0923 (0.1475) n= 524	-0.7714*** (0.3214) n= 272	-0.3666* (0.2117) n= 847	-0.5128** (0.2346) n= 717	-0.0251 (0.0922) n= 106
Evasão Escolar												
Idade 6-17 anos	0.0018 (0.0039) n=18459	0.0014 (0.0048) n= 14804	0.0009 (0.0105) n= 3174	-0.0651 (0.0625) n= 5757	-0.0275 (0.0322) n= 5305	-0.9833*** (0.0300) n= 327	-0.0035 (0.0053) n= 6100	-0.0061 (0.0084) n=4036	0.0005 (0.0074) n= 1897	0.0110 (0.0888) n= 6602	0.0107 (0.0108) n= 5463	0.0143 (0.0161) n= 950
Idade 6-14 anos	-0.0025 (0.0086) n=11972	0.0031 (0.0027) n= 9686	-0.1221 (0.1669) n=1994	-0.0962 (0.1374) n= 3793	0.0010 (0.0044) n= 3492	-0.0925 (0.0080) n=223	0.0014 (0.0028) n= 3833	0.0022 (0.0026) n= 2558	-0.0640 (0.1345) n=1179	0.0053 (0.0056) n= 4346	0.0057 (0.0063) n= 3636	0.0070 (0.0460) n= 592
Idade 15-17 anos	-0.0027 (0.0052) n= 4083	-0.0039 (0.0057) n= 3274	-0.0001 (0.0195) n= 669	-0.0472 (0.0539) n= 1278	-0.0662 (0.0738) n= 1168	-0.0579 (0.2286) n= 74	-0.0324 (0.0327) n= 1320	-0.0546 (0.0545) n= 868	0.0037 (0.1259) n= 404	0.0009 (0.0031) n= 1485	0.0011 (0.0029) n= 1238	-0.0146 (0.0524) n= 191

Fonte: Painel produzido a partir da AIBF I e AIBF II. Os valores entre parênteses são os erros padrão.

Nota: * Significativo ao nível de 10%

** Significativo ao nível de 5%

*** Significativo ao nível de 1%

Os resultados para matrícula escolar mostram que o PBF teve impacto significativo para as regiões Norte/Centro-Oeste e Nordeste. Para região Norte/Centro-Oeste, ao nível de significância de 1%, a probabilidade das crianças beneficiárias se matricularem na escola aumenta em 6,62% quando comparados a crianças que não são beneficiários do PBF. Para a zona urbana, a probabilidade de freqüentar a escola aumentou 5,39% e para a zona rural essa probabilidade foi de 70,84%, com nível de significância de 5 e 1 pontos percentuais, respectivamente. Pode-se observar também que o PBF aumenta a probabilidade das crianças beneficiárias de 6 a 14 anos residentes nessa região e na zona rural, de se matricularem na escola em 91,43%, em relação a crianças não beneficiárias para todo Brasil, de 2005 para 2009.

Esses resultados corroboram com a literatura, pois apesar da matrícula escolar das crianças brasileiras terem aumentado de um modo geral, a diferença entre as taxas de matrícula de crianças que vivem na área rural e crianças que vivem na área urbana continua apresentando disparidades, pois a área rural convive com uma oferta menor de escolas e uma infraestrutura pior quando comparada com a área urbana.

Para a região Nordeste, o PBF aumentou a probabilidade das crianças se matricularem na escola em aproximadamente 7%, para toda a região e para zona urbana. Para a região Sul/Sudeste e para as crianças mais velhas (15 a 17 anos) os impactos não foram estatisticamente significativos para matrícula escolar.

No entanto, para os indicadores educacionais de progressão e repetência escolar, esses resultados são significativos para a região Sul/Sudeste e para as crianças mais velhas, indicando que o PBF aumenta em 36,66% a probabilidade de essas crianças beneficiárias progredirem de série e diminuindo assim a probabilidade das mesmas repetirem de série. Este impacto também é positivo, aumentando a probabilidade das crianças de 15 a 17 anos em 51,28% de progredirem de série na área urbana, diminuindo também em 51,28% a probabilidade de repetência escolar. Um aumento de 77,14% na probabilidade das crianças de 15 a 17 anos, que participam do BVJ, pode ser observado para crianças da área rural residentes na região Nordeste.

Verifica-se também que para a evasão escolar, o PBF diminui em 98,33% a probabilidade das crianças entre 6 e 17 anos, residentes na área rural na região Norte/Centro-Oeste, evadirem a escola. Para as outras regiões o impacto não foi estatisticamente significativo.

Esses resultados indicam que progredir na escola tem um efeito mais significativo nos jovens beneficiários de 15 a 17 anos. O custo de se manter na escola parece maior para os jovens dessa faixa etária, principalmente os que residem em áreas rurais, que muitas vezes perdem o interesse pela escola ou não abrem mão de trabalhar. Isso indica que a criação do BVJ, fez com que esses jovens, que muitas vezes já haviam abandonado a escola ou precisavam trabalhar para complementar a renda da família, dedicassem mais tempo à escola, aumentando assim o seu progresso e os mantendo na escola.

6. CONSIDERAÇÕES FINAIS

O objetivo deste estudo foi avaliar o impacto do PBF sobre os indicadores educacionais das crianças beneficiárias de 6 a 17 anos. Os impactos do benefício foram investigados na matrícula escolar, que é uma condicionalidade do programa, e também na progressão, repetência e evasão escolar. Para as análises foi construído um painel de dados longitudinais inédito, a partir da AIBF I e AIBF II, de 2005 e 2009. Além disso, as análises consideram o desenho amostral, que ao ser ignorado, pode gerar viés nas estimativas e produzir resultados incorretos.

Quanto aos efeitos do programa sobre a matrícula escolar, os resultados mostraram que o PBF aumenta a probabilidade das crianças beneficiárias de se matricularem nas regiões Norte/Centro-Oeste e Nordeste, especialmente nas áreas rurais. Os efeitos marginais indicaram que a probabilidade das crianças do grupo de tratamento, residentes na região Norte/Centro-Oeste, de se matricularem na escola aumentou aproximadamente 91,43%.

Os resultados também mostraram que o programa impactou positivamente na progressão escolar para crianças de 15 a 17 anos residentes na área rural do Nordeste e na área urbana da região

Sul/Sudeste. Na área rural da região Norte/Centro-Oeste o PBF diminuiu em aproximadamente 98% a probabilidade de evasão escolar para as crianças de todas as idades. No entanto, nenhum resultado foi estatisticamente significativo quando a amostra foi separada por gênero.

Os resultados como um todo mostram que o PBF parece estar aumentando o atendimento escolar nas áreas rurais e na região Norte/Centro-Oeste do país. Nas áreas urbanas o PBF demonstra não estar cumprindo com seu objetivo de aumentar acumulação de capital humano entre as crianças mais pobres, reduzindo assim a transmissão intergeracional da pobreza.

Uma pesquisa realizada com beneficiários do PBF, de Sjoberg e Nilsson (2013) conclui que o benefício é de grande ajuda para as famílias e que a maior parte do subsídio é gasto em alimentação. No entanto, os entrevistados para essa pesquisa acrescentaram que as crianças teriam sido matriculadas na escola, mesmo que não recebessem o benefício do BF.

Assim, apesar de o PBF não ser uma política educacional e, portanto, não ter a função de melhorar a qualidade da educação, mudanças significativas no sistema de ensino fazem-se necessária para quebrar o ciclo intergeracional da pobreza via educação. Observa-se por parte do governo tentativas de se construir uma rede de proteção social em torno das famílias mais vulneráveis. Entretanto, essa iniciativa ainda não tem o alcance desejável.

BIBLIOGRAFIA

AGUILAR, Ernesto Yáñez. El impacto del Bono Juancito Pinto. Um análisis a partir de micro simulaciones. 2010.

AHMED, Akhter *et al.* Impact evaluation of the conditional cash transfer program in Turkey: Final report. International Food Policy Research Institute, Washington, DC, 2007.

BEHRMAN, Jere R.; PARKER, Susan W. Impacts of conditional cash transfer programs in education. Conditional cash transfers in Latin America, Baltimore, The John Hopkins University Press/International Food Policy Research Institute, p. 191-211, 2010.

CEDEPLAR. Descrição da Pesquisa AIBF – O Processo Amostral CEDEPLAR (2006). Projeto de avaliação do impacto do bolsa família [Impact assessment of Bolsa Família]. Final analytical report. Belo Horizonte, CEDEPLAR, 2005.

CURRALERO, C.B *et al.* AS CONDICIONALIDADES DO PROGRAMA BOLSA FAMÍLIA. In.CASTRO, J.A. MODESTO, L. (org.). Bolsa família 2003-2010: avanços e desafios. Brasília: IPEA, 2010. p. 151-173.

DE BRAUW, Alan *et al.* The impact of BolsaFamilia on child, maternal, and household welfare. IFPRI. Washington, DC,2012.

DE BRAUW, Alan *et al.* The Impact of Bolsa Família on Schooling. World Development, v. 70, p. 303-316, 2015.

FILHO, Walmir dos Reis Miranda. Avaliação do Impacto de Políticas de Transferência de Renda a partir de Dados Amostrais Complexos. Dissertação (Mestrado em Estatística) – Universidade Federal de Pernambuco, Pernambuco, 2017.

FILMER, Deon; SCHADY, Norbert. The Medium-Term Effects of Scholarships in a Low-Income Country. Journal of Human Resources, v. 49, n. 3, p. 663-694, 2014.

FISZBEIN, Ariel; SCHADY, Norbert Rüdiger; FERREIRA, Francisco HG. Conditional cash transfers: reducing present and future poverty. World Bank Publications, 2009.

GLEWWE, Paul; KASSOUF, Ana Lúcia. O impacto do Programa Bolsa Família no total de matrículas do ensino fundamental, taxas de abandono e aprovação. Anais do Encontro Nacional de Economia, v. 36, 2008.

- GLEWWE, Paul; KASSOUF, Ana Lucia. The impact of the Bolsa Escola/Familia conditional cash transfer program on enrollment, dropout rates and grade promotion in Brazil. *Journal of development Economics*, v. 97, n. 2, p. 505-517, 2012.
- KARACA-MANDIC, Pinar; NORTON, Edward C.; DOWD, Bryan. Interaction terms in nonlinear models. *Health services research*, v. 47, n. 1pt1, p. 255-274, 2012.
- KISH, L. *Survey Sampling*. New York: John Wiley & Sons, 1965.
- LEVY, Santiago. *Progress against poverty: sustaining Mexico's Progresa- Oportunidades program*. Brookings Institution Press, 2007.
- MALUCCIO, John A. *et al.* The impact of improving nutrition during early childhood on education among guatemalan adults*. *The Economic Journal*, v. 119, n. 537, p. 734-763, 2009.
- Ministério do Desenvolvimento Social e Combate à Fome (MDS). *Avaliação de Impacto do Programa Bolsa Família. - 1ª Rodada (AIBF I). Sumário Executivo*. Brasília, jun. 2007.
- Ministério do Desenvolvimento Social e Combate à Fome. (2017). Disponível em <http://mds.gov.br/assuntos/bolsa-familia>. Acesso em 13/07/2017.
- OLIVEIRA, Dalila Andrade; DUARTE, Adriana. Política educacional como política social: uma nova regulação da pobreza. *Perspectiva*, v. 23, n. 2, p. 279-301, 2005.
- PUHANI, Patrick A. The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters*, v. 115, n. 1, p. 85-87, 2012.
- REIMERS, Fernando; DA SILVA, Carol DeShano; TREVINO, Ernesto. *Where is the " education" in conditional cash transfers in education?*. Montreal: UNESCO Institute for Statistics, 2006.
- SCHADY, Norbert *et al.* Cash Transfers, Conditions, and School enrollment in Ecuador. *JOURNAL OF LACEA ECONOMIA*, 2008.
- SCHAFFLAND, Elke. *Conditional Cash Transfers in Brazil: Treatment Evaluation of the 'Bolsa Família' Program on Education*. Courant Research Centre: Poverty, Equity and Growth-Discussion Papers, 2012.
- SCHULTZ, T. Paul. School subsidies for the poor: evaluating the Mexican Progresa poverty program. *Journal of development Economics*, v. 74, n. 1, p. 199-250, 2004.
- SCHWARZMAN, Simon. Bolsa família: Mitos e realidades. *Interesse Nacional*, v. 2, n. 7, p. 20-28, 2009.
- SIMOES, Armando Amorim; SABATES, Ricardo. The Contribution of Bolsa Família to the educational achievement of economically disadvantaged children in Brazil. *International Journal of Educational Development*, v. 39, p. 141-156, 2012.
- SJÖBERG, Karin; NILSSON, Hanna. *An Evaluation of the Impacts of Bolsa Família on Schooling*. 2013.
- SKINNER, C. J.; HOLT, D. e SMITH, T. M. F. *Analysis of Complex Surveys*. New York: John Wiley & Sons, 1989.
- SOARES, Sergei; SÁTYRO, Natália. *O Programa Bolsa Família: desenho institucional, impactos e possibilidades futuras*. Texto para Discussão, Instituto de Pesquisa Econômica Aplicada (IPEA), 2009.
- WOOLDRIDGE, J. 2007. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.