

Impactos do Programa Bolsa-Família Sobre os Indicadores Educacionais

Rosana Ribeiro

Professora Associada do Instituto de Economia da Universidade Federal de Uberlândia (IE/UFU), Brasil

Maria Cristina Cacciamali

Professora Titular e Livre Docente da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA/USP), Brasil

Resumo

O artigo investiga os impactos do programa Bolsa-Família sobre a frequência escolar e a defasagem idade-série. Os dados são provenientes da PNAD (Pesquisa Nacional por Amostra de Domicílios) do ano de 2006 e a estratégia empírica se baseia no pareamento pelo escore de propensão. Os testes de diferença das médias desses indicadores educacionais não revelaram diferenças significativas entre as crianças e adolescentes dos grupos de tratamento e de controle. Noutras palavras, o programa mostrou-se inócuo em relação aos indicadores educacionais estudados.

Palavras-chave: Bolsa-Família, Brasil, Indicadores Educacionais, Pobreza

Classificação JEL: I28

Abstract

The article investigates the impact of the social program Bolsa-Família (Family Scholarship Program) on school attendance and the age-grade lag. The data are provenient from the PNAD (Pesquisa Nacional por Amostra de Domicílios – National Household Research) of the year 2006 and the empirical strategy is based on propensity score matching. The differential test of the mean education indicators presented no significant differences between children and adolescents of control and treatment groups. It can therefore be concluded that the program was innocuous in relation to the education indicators that were investigated.

* Recebido em maio de 2010, aprovado em novembro de 2012.
E-mail addresses: rosanaribeiro@prove.ufu.br, cciamali@uol.com.br

1. Introdução

O fenômeno da pobreza, mensurado a partir do conceito de pobreza absoluta ou pobreza multidimensional, ainda persiste no Brasil, embora seja inegável que no período recente a pobreza absoluta se reduziu entre os brasileiros. Na perspectiva dos defensores da teoria do capital humano, os reduzidos patamares desse capital – constituído pelas habilidades naturais e adquiridas – são os responsáveis pelos índices de pobreza absoluta. As habilidades adquiridas resultariam de decisões “voluntárias” dos indivíduos em relação à permanência na escola ou da opção por qualquer outro tipo de formação (exemplos: formação profissional, treinamento do local de trabalho) que possa melhorar esse capital e seja objeto de recompensa monetária no mercado de trabalho. Dentre as habilidades adquiridas, os teóricos do capital humano conferem maior atenção à permanência na escola. Segundo esses teóricos, a pobreza resulta de uma baixa acumulação de capital humano na medida em que o indivíduo portador desse baixo capital tem reduzida produtividade no trabalho e, conseqüentemente, baixos rendimentos. Cabe ponderar que é razoável supor que um trabalhador com mais tempo de treinamento produza mais, embora a associação entre nível de escolaridade e produtividade, para todos os postos de trabalho, seja no mínimo obscura. O estudo de Weiss (1988), sobre o desempenho de um conjunto de trabalhadores que montavam equipamentos, observou que cada ano de educação secundária aumentava a produção do trabalhador em somente 1,3%. Em geral, as pessoas mais escolarizadas têm maior probabilidade de serem selecionadas pelas empresas de maior produtividade. Nesse caso se trata também de um fenômeno de demanda e não somente de oferta.

Todavia, os adeptos da teoria do capital humano geralmente defendem uma estreita relação entre escolaridade, produtividade e salários. Sob essa ótica, o baixo capital humano se perpetua de geração em geração e origina, portanto, a armadilha da pobreza. As famílias pobres necessitam dos rendimentos de todos os membros que estejam aptos ao trabalho, assim, tão logo quanto possível, suas crianças são compelidas para o trabalho. O ingresso precoce no mercado de trabalho pode levar ao abandono da escola e ao comprometimento do capital humano dessas crianças. Na fase adulta, essas pessoas receberiam, então, baixos rendimentos e formariam novas famílias pobres.

É nesse contexto que o programa de transferência de renda condicionada implementado no Brasil busca superar a pobreza no longo prazo, por meio de uma assistência financeira às famílias carentes, no intuito de compensar a renda proveniente do trabalho infantil e, com isso, aumentar a frequência das crianças à escola. O objetivo desses programas é contornar a situação de pobreza absoluta por meio da melhora do capital humano das crianças e jovens pobres. Desse modo, a melhora em indicadores como permanência na escola e defasagem idade-série dos beneficiários desses programas deve ser alcançado, a fim de que o seu objetivo no longo prazo seja atingido.

Posto isso, o intuito deste estudo é analisar o impacto do Programa Bolsa-Família,

criado em 2004, sobre os indicadores educacionais, sobretudo na proporção de crianças e adolescentes que deixaram de comparecer às aulas em determinado período e na razão de crianças com defasagem idade-série. O atraso escolar é decorrente da entrada tardia na escola, da repetência ou da evasão escolar. Uma melhora no indicador de defasagem idade-série denotaria, portanto, alterações nesses determinantes. Assume-se aqui o pressuposto de que a ajuda financeira concedida pelo programa melhoraria os indicadores educacionais eleitos, na medida em que reduz, sobretudo, os custos indiretos¹ da educação para as famílias pobres. O programa poderia minimizar a defasagem idade-série estimulando a entrada da criança na escola na idade prevista pela legislação, reduzindo a evasão escolar, além do fato de que a frequência às aulas, exigida pelo programa, contribuiria também para melhorar a performance escolar das crianças e, com isso, a passagem para outra série escolar, diminuído, portanto, o atraso escolar.

Cabe ressaltar que, neste estudo, a investigação dos efeitos do programa Bolsa-Família sobre os indicadores educacionais das crianças e jovens beneficiárias não se limita às pessoas que se inserem na família na condição de filhos na medida em que incluímos também as crianças e adolescentes inseridas na condição de outros parentes ou agregados, além do que consideramos também o limite estabelecido pelo programa em relação ao número de crianças e/ou adolescentes que poderiam receber o auxílio monetário. Ou seja, no máximo 3 pessoas com idade entre 7 e 15 anos seriam candidatas ao benefício em 2006. Aqui parte-se do pressuposto que os adultos que exercem papéis-chaves nas famílias (chefe e/ou cônjuge) assumem perante todas as crianças e adolescentes funções semelhantes aquelas exercidas pelos pais. Além do que os impactos do programa Bolsa-Família sobre os indicadores educacionais eleitos são estudados em separados para as crianças e adolescentes da zona rural e urbana porque existem especificidades nas dificuldades encontradas em cada uma dessas áreas no que se refere ao acesso escolar.

Os microdados utilizados são provenientes dos questionários suplementares da Pesquisa Nacional por Amostra de Domicílios (PNAD) do ano de 2006 e a metodologia se baseia no pareamento pelo escore de propensão, que permite a obtenção dos grupos de tratamento (beneficiários) e de controle (não beneficiários), reduzindo o viés de seleção que estaria presente numa comparação direta. O software utilizado é *Stata 10*.

O artigo é composto de seis seções, além desta introdução. A primeira compreende uma análise dos indicadores educacionais brasileiros no período recente enquanto que a seção seguinte apresenta uma síntese das avaliações dos programas de transferências de renda condicionados já realizadas no continente Americano. A terceira e quarta seções tratam da estratégia empírica utilizada e da metodologia

¹ Segundo os defensores da teoria do capital humano, a decisão de investir em educação somente se realiza se os benefícios superarem os custos, que se dividem em diretos e indiretos. Os primeiros custos são medidos pelos gastos com cada ano de estudo, como o pagamento de mensalidade escolar e livros, enquanto os custos indiretos correspondem aos salários não recebidos. Os benefícios, por sua vez, são dados pelo diferencial de salários obtido pelo ano extra de estudo. No Brasil, os estados e municípios ofertam educação pública e gratuita no ciclo básico, assim, minimizam os custos diretos.

de uso da PNAD, e a quinta é dedicada à interpretação dos resultados encontrados. Por último, as notas conclusivas.

2. Indicadores do Sistema Educacional Brasileiro

Atualmente, o sistema educacional brasileiro é composto por dois níveis: Educação Básica e Educação Superior. No que se refere à educação básica, verifica-se a seguinte segmentação: Educação Infantil, Ensino Fundamental e Médio.

O artigo 208 da Constituição Federal do Brasil estabelece que o Estado tem o dever de garantir a educação básica obrigatória e gratuita dos 4 (quatro) aos 17 (dezessete) anos de idade e assegurar inclusive sua oferta gratuita para todos aqueles que não tiveram acesso na idade própria. Por outro lado, uma nova legislação, Lei nº 11.274, sancionada em 06/02/2006, estendeu a duração do ensino fundamental de oito para nove anos. Os dados deste artigo se referem ao período em que vigorava a antiga legislação educacional, segundo a qual uma criança que tivesse 7 anos de idade até o início do ano letivo deveria cursar a primeira série do ensino fundamental, além do que essa legislação considerava também que esse nível de ensino teria duração de oito anos. Ademais, optou-se por analisar os impactos do Bolsa-Família sobre a defasagem de idade-série para aquelas crianças e jovens com idade de 07 a 15 de idade porque no ano de 2006 o programa condicionava a presença dessas pessoas na escola.

Nas últimas décadas, uma das características do sistema educacional brasileiro que merece atenção se refere a sua rápida expansão em todos os níveis de escolaridade. A taxa de atendimento² para crianças e adolescentes passou de 74% em 1981 para 93% em 1998 e, finalmente, no ano de 2008 atingiu 97%. Em outras palavras, registra-se a universalização do acesso ao ensino fundamental no país. A taxa de atendimento se elevou nesse período para todas as pessoas dessa coorte etária.

Segundo microdados da PNAD, no ano de 2008 a taxa de atendimento se encontrava em patamares superiores a 90% para as crianças e adolescentes de 7 a 14 anos de idade, moradores da zona rural ou urbana, integrantes de famílias situadas entre os 25% mais pobres da população ou inseridos em famílias cuja renda familiar per capita se localizava entre os 25% mais ricos. Essa taxa não atingia patamares iguais ou superiores a 90% somente para os jovens de 15 anos de idade residentes na zona rural ou inseridos em famílias pobres.

A expansão na oferta de escolas públicas e gratuitas, as mudanças socioeconômicas e espaciais das famílias, bem como transformações demográficas contribuíram para melhor cobertura do sistema escolar no Brasil. Dentre as alterações nas características familiares, destacam-se suas migrações para cidades, que por sua vez oferecem maiores oportunidades educacionais, além do fato de que no período recente nota-se inclusive uma redução da pobreza no país. Por

² Essa taxa corresponde ao número de alunos na escola em determinada faixa etária em relação à população na mesma faixa etária.

outro lado, a queda na fecundidade das mulheres³ e, portanto, a redução do tamanho das famílias provoca uma diminuição na demanda por vagas no ensino fundamental, sobretudo nas primeiras séries. Essa transformação demográfica facilita o desempenho do sistema escolar, já que a expansão populacional se constitui em um desafio para os formuladores de políticas educacionais. De acordo com Silva (2003), o menor número de filhos nas famílias também favorece a entrada e a permanência das crianças na vida escolar na medida em que reduz os gastos com educação das crianças e jovens por parte das famílias e eleva o tempo que os adultos podem dedicar a cada filho.

Além da frequência e permanência escolar é desejável que as crianças e jovens não apresentem defasagem idade-série no intuito de assegurar a conclusão do ciclo escolar básico na idade preconizada. No presente estudo, a estimativa da defasagem idade-série se inspirou nos procedimentos adotados por Machado e Gonzaga (2007), segundo os quais as crianças com 07 anos de idade no início do ano letivo deveriam cursar a antiga primeira série do ensino fundamental. Caso a criança realize as progressões escolares seguintes, não haverá defasagem escolar. Para o estudante (i) com mais de 07 anos, o cálculo da defasagem idade-série em determinado período da vida escolar (τ) pode ser expresso pelas fórmulas 1 e 2, a saber:

$$D_{i,\tau} = 0 \text{ se } \frac{EDUC_{i,\tau}}{Idade_{i,\tau} - 7} \geq 1 \quad (1)$$

$$D_{i,\tau} = 0 \text{ se } \frac{EDUC_{i,\tau}}{Idade_{i,\tau} - 7} < 1 \quad (2)$$

O termo $EDUC_{i,\tau}$ corresponde aos anos de estudos do aluno (i) no período (τ) e $Idade_{i,\tau}$ se refere à idade da criança ou adolescente (i) no início do ano letivo. Se essas pessoas estiverem em situação de defasagem escolar, então $D_{i,\tau} = 0$, e em caso contrário $D_{i,\tau} = 1$.

Como se vê nas Tabelas 1 e 2, a defasagem idade-série, estimada a partir das fórmulas anteriores, aumenta para as pessoas com idades mais elevadas. Acompanhando-se as diversas idades dos alunos no ensino fundamental, observa-se que as dificuldades de progressão escolar são cumulativas. Notam-se também diferenças no indicador de atraso escolar segundo o local de residência e renda familiar per capita. As crianças residentes na zona rural e aquelas inseridas em famílias situadas no primeiro quintil apresentam os maiores percentuais de defasagem idade-série, enquanto as crianças que moram no meio urbano e/ou que vivem em famílias do último quintil apresentam os menores percentuais.

Transcorridos dez anos, a proporção total de crianças e adolescentes com defasagem escolar se reduziu significativamente em todos os recortes. Contudo, as maiores diminuições nesse indicador se verificou entre pessoas residentes da zona rural e aqueles inseridos em família de baixos rendimentos. Alguns fatores podem explicar essa melhoria geral no atraso escolar como a expansão na oferta de

³ No Brasil, a taxa de fecundidade (número de filhos por mulher ao fim do período reprodutivo) caiu de 2,8 (1991) para 1,8 (2006).

Tabela 1

Distribuição das crianças e adolescentes com defasagem idade-série segundo zona de moradia e quintil de renda-1998-(%)

Idade escolar	Proporção de defasagem idade-série				
	global	residentes na zona urbana	residentes na zona rural	25% mais pobres	25% mais ricos
7	4	3	9	9	0,4
8	25	18	44	43	4
9	33	27	55	56	7
10	40	33	64	67	11
11	48	41	71	75	14
12	56	49	79	82	20
13	60	54	82	85	23
14	64	58	83	88	29
15	69	64	87	90	35
Total	45	39	64	65	17

Fonte: PNAD/IBGE, 1998. (Elaboração própria).

escolas públicas, sobretudo na zona rural; a ampliação da pré-escola e das classes de alfabetização; adoção de programa de aceleração de aprendizagem (sistemas de ciclos em substituição ao sistema seriado por ano ou promoção automática) em algumas unidades da federação; condições sociais mais favoráveis usufruídas pelas famílias; diminuição do trabalho infanto-juvenil, redução no percentual de crianças com entrada tardia na escola; diminuição da evasão escolar etc.

Em resumo, o país assiste à universalização do acesso de crianças e adolescentes à escola, entretanto o atraso escolar ainda se mostra elevado no país e permanece como objeto de preocupação para os responsáveis pelas políticas educacionais.

3. Avaliação dos Impactos dos Programas de Transferência de Renda Condicionados nos Indicadores Educacionais

Na concepção dos programas de transferência de renda condicionados, o círculo vicioso da pobreza no longo prazo seria quebrado por meio de uma assistência monetária proporcionada às famílias pobres no intuito de eliminar o trabalho infantil e elevar a frequência escolar enquanto no curto prazo a pobreza seria superada por meio da elevação dos rendimentos da família via o auxílio monetário. Esses programas se difundiram nos países subdesenvolvidos, sobretudo, após os anos 1990.

Tabela 2

Distribuição das crianças e adolescentes com defasagem idade-série segundo zona de moradia e quintil de renda-2008-(%)

Idade escolar	Proporção de defasagem idade-série				
	global	residentes na zona urbana	residentes na zona rural	25% mais pobres	25% mais ricos
7	2	2	2	2	1
8	22	21	28	29	12
9	28	25	40	38	14
10	31	27	45	43	11
11	34	31	49	47	14
12	40	36	55	55	15
13	44	40	57	61	17
14	45	42	59	63	18
15	46	42	64	66	22
Total	33	30	45	44	14

Fonte: PNAD/IBGE, 1998. (Elaboração própria).

Na América Latina, os programas de transferência de renda condicionados foram implementados inicialmente no Chile (Chile Solidario),⁴ México (Progresas)⁵ e Brasil (Bolsa-Escola e Bolsa-Família). Em 1997, o governo mexicano implementou o Progresas, cujo objetivo geral seria melhorar as condições de educação, saúde e nutrição das famílias pobres por meio de serviços na área educacional e de saúde, bem como via o fornecimento de assistência monetária e de suplementos nutricionais. O programa contemplava três objetivos específicos:

- a) encorajar a educação das crianças por meio do aumento das matrículas e elevação da frequência escolar, além de promover entre os pais a criação ou a consolidação da consciência acerca da importância da educação para seus filhos, e paralelamente ampliar ações para melhorar a qualidade da educação e assegurar que a qualidade das escolas não seria reduzida com a elevação da frequência escolar;
- b) ampliar a qualidade dos serviços de saúde para todos os membros da família, assim como divulgar as medidas preventivas de cuidados com a saúde e a

⁴ Esse programa foi implementado em 2002 e concedia apoio à família via assistente social, subsídios monetários aos beneficiários e acesso prioritário a outros programas de proteção social.

⁵ Em 1997, o governo mexicano criou o Progresas que deu origem mais adiante ao Programa Oportunidades, que fornece \$250 pesos por idoso, além de \$189 pesos para a compra de alimentos e bolsa de estudo no valor de \$120 pesos para as crianças que cursam o ensino primário enquanto para os jovens que cursam décima segunda série recebem \$760 pesos. O novo programa ampliou os beneficiários na medida em que incluiu também os residentes da área urbana (Soares et alii 2007).

nutrição das pessoas;

- c) realizar transferência monetária de \$189 pesos por mês e de suplementos para as famílias pobres, no intuito de melhorar o consumo e o padrão alimentar dessas famílias (Berhrman et alii 2001).

As crianças e jovens das famílias beneficiárias deveriam frequentar no mínimo 85% das aulas e realizarem visitas aos postos de saúde. Nota-se que o Progesa foi desenhado para desenvolver ações que ultrapassam a assistência financeira às famílias. No campo educacional, segundo Berhrman et alii (2001), existe uma atenção dirigida à conscientização das famílias sobre a importância da educação de seus filhos. Ademais, o benefício monetário se eleva com a idade das crianças e assume maior valor para as meninas, porque, segundo Soares et alii (2007), existe maior evasão escolar dessas crianças na zona rural mexicana.

Berhrman et alii (2001) avaliam o impacto do Progesa sobre a frequência escolar, a repetência, a evasão e a taxa de retorno escolar entre as crianças que abandonaram a escola nos primeiros dois anos de operação desse programa (1998/1999). Os beneficiários do programa apresentaram maiores taxas de frequência às aulas, menor índice de repetência, menor nível de evasão e maior taxa de retorno entre as crianças que abandonaram a escola, em relação às crianças do grupo de controle. Um resultado ressaltado pelos autores se refere à notável redução de evasão escolar entre os beneficiários aptos a transitarem do ciclo escolar básico para o ciclo secundário. Os impactos do Progesa para meninos e meninas são similares na etapa do ensino primário, porém o programa apresenta melhores resultados no combate à evasão entre o ciclo básico e secundário para os meninos.

Essa modalidade pública de programa de enfrentamento à pobreza se constitui em um instrumento utilizado por distintos governos como Honduras (PRAF- Programa de Asiganción Familiar-1990), Colômbia (Famílias em Acción-2000), Nicarágua (Mi Familia-2007) e Costa Rica (Superémonos-2000). Nesse último país, o programa se baseia na transferência condicionada de cupons de alimentação que no entanto apresenta condicionalidades semelhantes àquelas dos programas de transferência de renda.

O Superémonos fornece cupons de alimentação para todas as famílias pobres, porém estabelece que todas as crianças e jovens na faixa de 6 a 18 anos devem frequentar a escola. Em 2001, o valor do vale alimentação correspondia a 30 dólares por mês e atingia cerca de 12.200 famílias. Outra condicionalidade do programa se refere à exigência do gasto dos cupons apenas com produtos alimentícios. Os beneficiários do Superémonos passam por um teste de renda e são elegíveis dependendo dos seguintes critérios: ocupação do chefe da família, renda familiar e nível educacional da pessoa de referência do domicílio.

Duryea e Morrison (2004) avaliam o impacto desse programa sobre a frequência escolar, a transição escolar e o trabalho infantil. Os resultados revelaram que o programa teve um impacto positivo e significativo sobre a frequência escolar, porém não teve efeito sobre a progressão escolar e o trabalho infantil. O programa implementado na Costa Rica não é efetivamente de transferência de renda, mas tem condicionalidades relativas à educação das crianças e jovens. Esse

programa, diferentemente do Progres, não busca conscientizar as famílias acerca da importância da educação de seus filhos. Essa última distinção pode ser a origem do tímido efeito do programa na Costa Rica sobre o indicador de transição escolar.

No Brasil temos, desde o ano de 2004, o Programa Bolsa-Família. Os componentes da família são inscritos no programa após a decisão de algum membro do núcleo familiar e, portanto, não é um programa dirigido ao indivíduo. A renda é o parâmetro para decidir quais famílias poderão ser beneficiárias do programa, além do que as famílias são diferenciadas em dois cortes de renda. Em setembro de 2006, as famílias com renda per capita inferior a R\$ 50,00 por mês estavam aptas a receber o valor de R\$ 50,00 por mês e aquelas famílias que tinham filhos com até 15 anos poderiam receber mais R\$ 15 por criança, porém no máximo três benefícios seriam pagos. O valor máximo pago para essas famílias seria R\$ 95,00. Por outro lado, as famílias com renda entre R\$ 50,00 e R\$ 100,00 somente poderiam receber os benefícios relativos às crianças. Nesse caso, as famílias receberiam no máximo R\$ 45,00. Em contrapartida, as crianças e adolescentes devem frequentar 85% das aulas.

Pesquisadores brasileiros analisaram mais intensamente os impactos do Bolsa-Escola e Bolsa-Família sobre o trabalho infantil (Ferro e Kassouf 2005; Cacciamali et alii 2009) e os gastos de consumo (Resende 2006). No período mais recente, Romero e Hermeto (2009) avaliaram de forma pioneira os efeitos desse programa sobre os indicadores educacionais, quais sejam, a proporção de crianças no domicílio que deixaram de ir à escola no último mês; a razão de crianças no domicílio que evadiram do sistema de ensino entre 2004 e 2005; a proporção de crianças no domicílio que foram aprovadas entre 2004 e 2005; a razão de crianças que declararam somente estudar em relação às que trabalham e estudam, em comparação com aquelas que apenas trabalham e em relação àquelas que não trabalham e não estudam; e a proporção de crianças reprovadas entre 2004 e 2005. Esses indicadores foram desagregados segundo o sexo e recortes geográficos. Cabe observar que esses indicadores têm repercussão sobre a defasagem idade-série, embora os autores não investiguem o atraso escolar.

Os dados utilizados são oriundos da Pesquisa de Avaliação de Impacto do Programa Bolsa-Família (AIBF) realizada em 2005 e do Cadastro Único para programas sociais. Os autores utilizaram a regressão descontínua em suas estimações. A amostra se restringiu às famílias com crianças e jovens entre 7 a 14 anos de idade, sendo que na AIBF o número de domicílios do grupo de tratamento foi de 3.248 e no grupo de controle 2.568. No Cadastro Único, o número de domicílios do grupo de tratamento foi de 3.988, e no grupo de controle 2.955.

Em relação ao indicador de evasão escolar, a estimativa por meio da regressão descontínua revelou que os maiores beneficiários do programa são meninas residentes em todo o Brasil e cujo rendimento familiar é inferior a R\$ 50 e R\$ 100; crianças e adolescentes nordestinos que vivem em lares cuja renda é inferior a R\$ 100 e aqueles que residem nas regiões Norte/Centro-Oeste com renda familiar menor que R\$ 50 tiveram menor proporção de abandono em relação ao grupo de controle. Focalizando a proporção de aprovados, os estudantes nordestinos do

sexo masculino do grupo de tratamento, inseridos em famílias com renda inferior ou igual a R\$ 50, tiveram desempenho superior ao grupo de controle. No caso da repetência escolar, verifica-se menor proporção de reprovados entre as crianças e adolescentes nordestinos do grupo de tratamento, que vivem em famílias com renda per capita de no máximo R\$ 100, e os meninos nordestinos cujas famílias têm renda familiar inferior ou igual a R\$ 50. Romero e Hermeto (2009) também constataram que a razão de pessoas de 7 a 14 anos de idade que trabalham é maior no grupo de tratamento em relação ao grupo de controle. Para a variável que capta a ausência escolar, não se registraram divergências entre grupo de tratamento e de controle, em nenhum nível de desagregação. Todavia, os autores apontam que os resultados não significativos podem ser creditados à configuração do modelo de regressão descontínua, na medida em que as famílias que se posicionam no extremo ou com renda distante do ponto de corte não são contempladas. Desse modo, a pesquisa revelou que se considerarmos os dados relativos à totalidade de crianças e adolescentes do país, somente o indicador de evasão escolar para as meninas foi significativo e favorável aos beneficiários do programa.

Vale destacar que a amostra das regiões Norte e Centro-Oeste, oriunda da AIBF, compreende, nos grupos de tratamento e de controle, respectivamente 1.221 e 1.400 domicílios, enquanto no Cadastro Único esses números correspondem a 1.586 e 1.502. Para a região Nordeste, o número de domicílios provenientes da AIBF totaliza 1.116 e 1.565, no grupo de tratamento e de controle, respectivamente, ao passo que no caso do Cadastro Único esses números são 1.900 e 1.735. Essas amostras são reduzidas ainda mais segundo os recortes por gênero e faixa de rendimentos. Em suma, a amostra de domicílios utilizada em recortes com elevado grau de desagregação pode se revelar pequena, porém os autores não fornecem detalhamentos sobre a montagem da amostra utilizada na pesquisa. Ademais, os testes de significância consideram aparentemente o desenho amostral como aleatório.

4. Estratégias Empíricas

Em nosso estudo utilizamos o método do Pareamento pelo Escore de Propensão (PEP) no intuito de obtermos as crianças e adolescentes beneficiadas e não beneficiadas pelo Programa Bolsa-Família. As estimativas das diferenças de médias dos indicadores educacionais eleitos entre esses grupos não se baseou no comando *psmatch2* disponível no *Stata* porque nesse caso não podemos considerar o desenho amostral estratificado. Diante disso, realizaram-se alguns procedimentos com os dados da PNAD que são necessários para o cálculo dessas diferenças por meio do comando *lincom* que nos permite considerar o desenho amostral como estratificado. Vale observar que a realização de testes de diferenças de médias que considerem o desenho da PNAD como amostra aleatória simples tem estimativas de erros padrões com viés para baixo.

4.1. Pareamento pelo Escore de Propensão

Optou-se neste estudo pela metodologia do pareamento pelo escore de propensão, no intuito de analisar os efeitos do programa Bolsa-Família sobre a ausência escolar e o atraso escolar. O procedimento permite a obtenção de um grupo de controle com características observáveis semelhantes às aquelas existentes no grupo de tratamento. Após a obtenção desses grupos, se estimam as diferenças médias entre os grupos de tratamento e de controle em relação às variáveis de resposta (ausência escolar e atraso escolar).

Um dos problemas que emergem ao se construir o grupo de controle é o denominado viés de seleção, que resulta das diferenças que podem existir entre suas características observáveis e não observáveis e aquelas do grupo de tratamento. Nos estudos das ciências sociais não podemos observar o mesmo indivíduo antes e depois do acesso ao tratamento, portanto deve-se resolver o viés. Além disso, o programa Bolsa-Família não foi distribuído aleatoriamente, mas obedece a um processo de seleção na medida em que existem critérios de elegibilidade definidos. Caso a aleatoriedade fosse assegurada, o grupo de controle teria a mesma distribuição de características do grupo de tratamento. Posto esse desafio, a abordagem do pareamento pelo escore de propensão (*propensity matching*), desenvolvido por Rosenbaum e Rubin (1983), surge como uma possível solução para o problema do viés de seleção.

Vários estudiosos tentaram superar as dificuldades na montagem desses grupos, como Roy (1951) e Rubin (1974), que propuseram a substituição do grupo de controle pelos indivíduos que efetivamente não recebem o tratamento. Se introduzirmos uma variável dummy que estabeleça 1 para os participantes e 0 para os não participantes do programa – $Y_{i(1)}$ corresponde à variável de resultado para o indivíduo i do grupo de tratamento, enquanto $Y_{i(0)}$ representa a variável resultado do indivíduo i do grupo de controle – se torna possível, numa linguagem algébrica, expressar o resultado do tratamento⁶ (Y) sobre cada indivíduo (i) por meio da fórmula:

$$\pi_i = Y_{i(1)} - Y_{i(0)} \quad (3)$$

Por outro lado, a estimação média dos resultados de um programa ou benefício é representada pela equação:

$$\pi_i = E [Y_{i(1)}/D_i = 1] - E [Y_{i(0)}/D_i = 1] \quad (4)$$

A equação anterior fornece o efeito médio do tratamento sobre os beneficiários. Esse caminho exige que se observem as pessoas antes e após o tratamento, porém, conforme mencionado acima, em ciências não experimentais somente temos a informação do indivíduo a partir de sua incorporação ao grupo de tratamento. Uma saída, portanto, é a substituição desse grupo por pessoas que efetivamente não

⁶ Neste caso o tratamento se refere ao programa de transferência de renda.

participam do programa $E[Y_{i0}|D_i = 0]$. Nesse caso, a equação 5 se transformaria na seguinte fórmula:

$$\pi_i = E[Y_{i(1)}|D_i = 1] - E[Y_{i(0)}|D_i = 0] \quad (5)$$

Essa substituição fornece uma estimação do impacto do tratamento sobre o grupo de controle, mas com viés, na medida em que existem distinções entre $E[Y_{i(0)}|D_i = 0]$ e $E[Y_{i(0)}|D_i = 1]$. Nesse caso, o viés, portanto, se origina da utilização de um grupo que representaria aqueles beneficiários no período anterior a sua incorporação ao programa. Heckman et alii (1998) dividem o viés em três componentes. O primeiro resulta da falta de suporte comum entre beneficiários e não beneficiários, ou seja, existiriam divergências nas características observáveis, enquanto o segundo componente se deve às diferenças na distribuição dessas características entre o grupo de tratamento e de controle. Por último, o terceiro viés resultaria das divergências de resultados que seriam encontradas, ainda que as características observáveis sejam controladas. Nesse caso, o viés se deve à presença de variáveis não observadas que influenciam os impactos potenciais e a participação no programa.

No intuito de minimizar o viés, se introduz uma hipótese de identificação que estabelece que o processo de seleção se realize por meio das características observáveis dos indivíduos ou famílias, definidos por X . Em linguagem algébrica tem-se a equação:

$$E[Y_{i(0)} - Y_{i(1)}] = E[Y_{i(0)}|D_i = 0, X] - E[Y_{i(1)}|D_i = 1, X] \quad (6)$$

Ou seja, indivíduos ou famílias com características semelhantes têm igual probabilidade de serem alocados no grupo de tratamento ou de controle. A partir dessas considerações, se assume que os resultados potenciais independem da participação no programa, uma vez conhecidas as características observáveis, conforme expressa a igualdade seguinte:

$$E[Y_{i(0)}, Y_{i(1)}|D_i, X] \text{ e } E[Y_{i(0)}|D_i = 0, X] = E[Y_{i(1)}|D_i = 1, X] \quad (7)$$

As equações anteriores, na verdade, se baseiam numa hipótese relevante para o método do pareamento por escore de propensão, denominada de independência condicional (Caliendo e Kopeinig 2005).

O grupo de controle seria identificado a partir das características observadas. Contudo, esse procedimento esbarra em limites porque, ao se elevar o número de características consideradas, se torna mais difícil encontrar um grupo de controle com as mesmas características. Noutras palavras, uma das dificuldades da obtenção de indivíduos semelhantes se deve à diversidade de características que são consideradas, ou seja, existem dificuldades em lidar com a dimensionalidade envolvida nesse processo. Rosenbaum e Rubin (1983) sugerem o uso do escore de propensão para resolver tais dificuldades. Os autores definem o escore de propensão como a probabilidade de um indivíduo ser incluído num programa a partir de suas características. Ou seja:

$$P(X) = \text{Probabilidade}(D = 1|X) \quad (8)$$

O problema da multidimensionalidade se resolve, então, a partir da adoção de um escalar. Os autores sugerem que $P(X)$ deve substituir X , conforme equação 9:

$$E[Y_{i(0)} - Y_{i(1)}|D_i = 1, P(X)] = E[Y_{i(0)}|D_i = 0, P(X)] - E[Y_{i(1)}|D_i = 1, P(X)] \quad (9)$$

O pareamento pelo escore de propensão permite a eliminação dos dois componentes iniciais do viés. O método elimina, portanto, o viés que se origina das características observáveis, porém aquele oriundo das não observáveis não pode ser controlado. Assim, o pareamento minimiza, mas não elimina inteiramente o viés de seleção. Na verdade, o procedimento do pareamento pelo escore de propensão somente pode ser desenvolvido se duas hipóteses são adotadas.

O primeiro pressuposto, mencionado anteriormente, se refere à independência condicional que estabelece que o grupo de tratamento e os impactos potenciais do programa são independentes das variáveis de pré-tratamento, ou seja, os resultados potenciais independem da participação no programa dadas as características das variáveis observadas (X). Na linguagem algébrica:

$$Y_{(0)}, Y_{(1)} \amalg D|X \quad (10)$$

onde \amalg representa independência.

Ou seja, pressupõe-se que a seleção é baseada em características observáveis e que todas as variáveis capazes de influenciar a participação no programa são controladas pelo pesquisador. Evidentemente que essa hipótese é bastante forte.

Rosenbaum e Rubin (1983) assumem também que o tratamento e os impactos potenciais são independentes da probabilidade em receber o tratamento, dado o escore de propensão. Essa suposição poderia ser expressa como:

$$Y_{(0)}, Y_{(1)}|P(x) \quad (11)$$

Outra hipótese importante para o cálculo do escore de propensão é a do suporte comum. Essa suposição estabelece que pessoas do grupo tratamento com características X tenham correspondentes no grupo de controle. Em termos do escore de propensão, a hipótese estabelece que para cada probabilidade estimada para indivíduos do grupo de tratamento tem-se uma probabilidade semelhante no grupo de controle.

À medida que se assegure a existência dessas hipóteses, o impacto médio do tratamento (ATT) pode ser expresso pela equação:

$$ATT = E\{E[Y_{(i)1}|D_i = 1, p(x_i)] - E[Y_{(i)0}|D_i = 0, p(x_i)] | D_i = 1\} \quad (12)$$

Noutras palavras, o efeito médio do tratamento é obtido pela diferença entre o resultado médio do grupo de tratamento e do grupo de controle. A estimativa do escore de propensão não permite o cálculo do ATT , porque seria praticamente

impossível encontrar duas pessoas com o mesmo valor do escore de propensão, na medida em que essa variável é contínua. Alguns algoritmos desenvolvidos tentam solucionar essa dificuldade.

Nesse artigo utilizamos o método de pareamento pelo vizinho mais próximo, em que a escolha dos indivíduos do grupo de controle se baseia na proximidade de seus escores de propensão em relação aos escores das pessoas do grupo de tratamento. Outro algoritmo é o pareamento pelo vizinho mais próximo com caliper, que exige que o pareamento seja realizado em um nível máximo de tolerância em relação à distância do escore de propensão.⁷ Por outro lado, o procedimento denominado Mahalanobis se baseia na distância euclidiana padronizada e agrega também a informação acerca da covariância existente entre as variáveis. Além disso, a diferença entre as médias das variáveis de resultados foi estimada via bootstrapping, que corresponde a uma técnica de reamostragem. Esse procedimento realiza a replicação de inúmeras amostras selecionadas, com reposição da amostra original e mesmo tamanho n dessa amostra. Davidson e Mackinnon (2000) recomendam que a estimativa de intervalos de confiança requeira em torno de 2000 replicações. Uma vantagem da utilização do bootstrapping é que no caso de amostra complexa esse método procura imitar, para a geração das novas amostras, o mesmo procedimento adotado na construção da amostra da PNAD. Nesse caso, a partir da amostra original da PNAD são selecionados os PSUs (*primary sampling unit* – unidade primária de amostragem) com reposição e probabilidade proporcional ao tamanho (Neder 2008).

4.2. Metodologia de uso da PNAD nos Testes de Diferenças de Médias

O delineamento da amostra da PNAD segue um esquema misto, com pré-estratificação de grupos regionais de municípios e a constituição de conglomerados em múltiplas etapas. Para cada Unidade da Federação (UF), subdivide-se sua área em diversos estratos, que são agrupamentos de diversos municípios vizinhos. Além disso, os municípios são classificados em três áreas censitárias: região metropolitana, municípios autorrepresentativos e municípios não autorrepresentativos. No caso das regiões metropolitanas e dos municípios autorrepresentativos, a unidade primária de amostragem (PSU) é o setor censitário na medida em que todos os municípios integram a amostra e em seguida selecionam-se os domicílios.

Nos estratos que contêm os municípios não autorrepresentativos em cada UF, inicialmente são selecionados dois municípios com probabilidade proporcional ao seu tamanho (estimativa populacional baseada no último Censo Demográfico). Em seguida, diversos setores censitários são selecionados e por último selecionam-se os domicílios. A amostra da PNAD tem um delineamento misto, pois nas regiões metropolitanas e municípios autorrepresentativos, o processo de seleção

⁷ Para maiores detalhes, ver Caliendo e Kopeinig (2005). Em nosso caso, utiliza-se *caliper* de 0,1% e *trimming* de 5%.

é realizado em dois estágios ao passo que nos municípios não autorrepresentativos compreende três estágios. Esse tipo de amostragem, que reduz consideravelmente os custos operacionais, eleva substancialmente os erros probabilísticos em relação aos correspondentes a uma amostra aleatória simples, pois em cada UF as unidades domiciliares ficam concentradas em um conjunto mais restrito de áreas. Esse procedimento reduz a diversidade de informação captada da população e eleva a variância amostral dos estimadores utilizados.

Neste artigo, todas as estimativas de indicadores educacionais levaram em consideração essas características da amostra da PNAD, para isso utilizando duas variáveis que definem o desenho da amostra: o estrato a que pertence o domicílio e a unidade primária. A partir desses procedimentos, o software *Stata* foi utilizado na estimativa dos indicadores, uma vez que, por meio de rotinas desse software, pode-se obter o erro padrão dos indicadores e, portanto, o grau de precisão das estimativas para os diversos cortes amostrais. Essas rotinas permitem, então, o cálculo dos testes estatísticos de diferenças dos indicadores analisados, considerando-se o delineamento de amostra complexa.

A elaboração das rotinas se baseia nos procedimentos desenvolvidos por Neder (2008), que enfrentou algumas dificuldades provenientes da existência de estratos com PSU único, na medida em que os métodos adotados pelo *Stata*⁸ exigem um mínimo de dois PSUs em cada estrato. Desse modo, optou-se por construir uma rotina para identificá-los e agregá-los aos estratos de maior número de observações, em cada UF, diluindo uma possível fonte de viés na estimativa da variância dos estimadores. A existência de estratos com PSU único se deve à criação de novos estratos referentes a novas unidades domiciliares, verificados na atividade anual de cadastramento realizada pelo IBGE. Em geral, o número de domicílios nesses estratos não é muito elevado em comparação ao conjunto da amostra. Após a adoção dos procedimentos necessários para que se considerasse a PNAD uma amostra estratificada, realizou-se os testes de diferenças de médias dos indicadores educacionais entre o grupo de tratamento e controle. Em resumo, esses procedimentos permitem que o desenho amostral seja considerado, além do peso amostral das famílias.

5. Dados e Variáveis

Os dados utilizados neste estudo são provenientes dos questionários suplementares da PNAD no ano de 2006. Um dos questionários se intitula “Características de acesso a algumas transferências de renda de programas sociais nos domicílios”. Segundo a metodologia da PNAD do ano de 2006, os quesitos dessa parte se concentraram em unidades domiciliares em vez de pessoas ou

⁸ O comando do *Stata* denominados *svy:mean*, que estima o valor de uma média (nesta pesquisa estimamos a média dos indicadores educacionais no grupo de tratamento e controle) utiliza em sua abordagem analítica o método da linearização de Taylor, baseado na conhecida decomposição matemática da fórmula do estimador e aplicada ao delineamento de amostragem por conglomerados.

grupos familiares, por razões operacionais. Vamos nos limitar aos domicílios que declararam que algum morador recebeu dinheiro do programa Bolsa-Família e que tem crianças entre 7 e 15 anos de idade. Além do que, delimitamos nossa amostra às famílias que têm no máximo 3 crianças e/ou jovens na faixa de 7 a 15 anos de idade, porque o programa estabelece o limite de três beneficiários. Essa restrição foi adotada como um procedimento de cautela, na tentativa de abarcamos efetivamente os beneficiários em idade escolar. Cabe observar que se optou pela eliminação dos domicílios que declararam receber auxílio do PETI (Programa de Erradicação do Trabalho Infantil), no intuito de evitar a fusão de programas distintos, embora ambos sejam dirigidos à população pobre e condicionem a presença das crianças na escola.

A PNAD do ano de 2006 tem também um suplemento dedicado às características complementares de educação e acesso à merenda escolar dos moradores de 0 a 17 anos de idade, que é útil na montagem das variáveis de resposta aqui investigadas.

Por outro lado, dados anteriores revelaram que as crianças e adolescentes residentes na área rural têm maior atraso escolar em relação àqueles que moram na zona urbana. Posto isto, investiga-se em separado o impacto do Bolsa-Família sobre os indicadores educacionais das crianças e adolescentes residentes na zona urbana e na zona rural. A partir do pareamento pelo método do escore de propensão via algoritmo do vizinho mais próximo, o número de domicílios dos grupos de tratamento e de controle na zona urbana atingiu, respectivamente, 7.935 e 15.765, enquanto os domicílios da área rural no grupo de tratamento são 3.145 e no grupo de controle 3.041. Na construção do grupo de controle nos limitamos as famílias cuja renda per capita familiar se revela igual ou inferior a R\$ 300,00.⁹ Adota-se esse procedimento com o propósito de melhorar a qualidade do pareamento na medida em que a renda familiar é o principal critério de seleção do programa.

As características das famílias é que determinam a possibilidade de inclusão no programa. Neste estudo, os modelos logits utilizados nos algoritmos de pareamentos levaram em conta essas características como nível de renda familiar per capita, escolaridade do responsável, tamanho da família etc. Cabe observar que nas amostras selecionadas em todos os domicílios somente existia uma única família.

As variáveis de resultados são a proporção de crianças e adolescentes que não compareceram à escola, em cada família, e a razão dessas pessoas com atraso escolar. Em cada família, a razão de crianças e adolescentes que deixaram de comparecer às aulas é obtida pela divisão entre as pessoas de 7 a 15 anos de idade que declararam que não compareceram pelo menos 1 dia à escola no período de 2 de agosto a 30 de setembro de 2006 e aqueles na mesma faixa etária que afirmaram frequentar regularmente a escola. O cálculo do atraso escolar se baseou nos procedimentos desenvolvidos por Machado e Gonzaga (2007), relatados em seção anterior. Identifica-se na família as pessoas com atraso escolar que têm entre

⁹ Esse recorte de renda permite que 95% das famílias da amostra dos beneficiários do Programa Bolsa-Família que tem filhos, parentes ou agregados de 7 a 15 anos de idade estejam inseridas no pareamento.

7 e 15 anos de idade. Em seguida, esse total é dividido pelo total de crianças e adolescentes na mesma faixa etária inseridos na família.

5.1. *Variáveis Independentes e Dependentes*

Para este estudo foram elaborados dois modelos logits, sendo um para o meio urbano e o outro para o meio rural. As variáveis independentes são dicotômicas e expressam as diferenças em termos de participação no programa Bolsa-Família. Uma delas foi codificada sob a forma $BolsaFamUrbano = 1$ se família residente no meio urbano e beneficiária do programa e $BolsaFamUrbano = 0$ se família residente no meio urbano e não beneficiária do programa. A outra variável independente é classificada $BolsaFamRural = 1$ se família residente no meio rural e beneficiária do programa e $BolsaFamRural = 0$ se família residente no meio rural e não beneficiária do programa.

As variáveis independentes são idênticas nos dois modelos logits e se referem aos parâmetros que influenciam a inclusão no programa Bolsa-Família. Em geral, os recursos familiares econômicos, culturais ou aqueles que se expressam por meio da relação entre os membros da família (capital social) se diferenciam segundo o estágio de pobreza e podem ser definidos. Além do que, existem distinções acerca das características individuais dos responsáveis pela família e do acesso a infraestrutura segundo o estágio de pobreza.

Em síntese, as variáveis independentes que refletem as características individuais são expressas por meio do sexo (*sexchefe*) e da raça (*raçachefe*). Para capturar a situação econômica das famílias, introduziu-se a variável renda familiar per capita (*yfam1qpc*), que é medida em termos de reais de setembro de 2006 (Quadro 1). Uma vez que investigamos o impacto de um programa de transferência de renda, optamos por deduzir da renda da família beneficiária aquela parcela oriunda do Bolsa-Família. Os procedimentos para essa dedução foram desenvolvidos a partir do estudo de Soares et alii (2006). O nível cultural é captado pela variável anos de estudo do chefe de família (*Anoestupai sum*), que é mensurada por meio dos anos de escolaridade. No intuito de apreender os aspectos da estrutura familiar, se inseriu a variável composição da família (*tipofam*).

A situação de pobreza também pode ser apreendida por meio do acesso aos serviços de infra-estrutura e de comunicação, pela posse de bens de consumo duráveis e características dos domicílios, bem como pela localização geográfica. O indicador de acesso do domicílio aos serviços de infraestrutura pode ser captado por meio das condições de acesso a água (*origemagua*) e destino do lixo (*destinolixo*). O grau de precariedade das condições do domicílio pode ser apreendido por meio do uso compartilhado do banheiro por diversas moradias (*usobanheiro*). Variáveis independentes relacionadas ao estado físico das moradias também se encontram no modelo, como o material das paredes (*matparede*) e o material de cobertura (*matcobertura*).

Quadro 1 – Variáveis independentes dos modelos logit utilizados nos pareamentos pelo escore de propensão

Variáveis	Descrição
yfamqlpc	Renda per capita familiar
Tipofam2	Variável dummy para família que tem somente pessoa de referência ¹⁰
Tipofam3	Variável dummy para outro tipo de arranjo familiar
regarecens2	Variável dummy para região Norte e área censitária metropolitana
regarecens3	Variável dummy para região Nordeste e área censitária metropolitana
regarecens4	Variável dummy para região Nordeste e interior
regarecens5	Variável dummy para região Sudeste e área censitária metropolitana
regarecens6	Variável dummy para região Sudeste e interior
regarecens7	Variável dummy para região Sul e área censitária metropolitana
regarecens8	Variável dummy para região Sul e interior
regarecens9	Variável dummy para região Centro Oeste e metrôpole
regarecens10	Variável dummy para Centro Oeste e área censitária interior ¹¹
Anoestupai	Anos de estudo do chefe da família
origemagua2	Variável dummy para acesso inadequado ¹²
origemagua10	Variável dummy para informação sobre condições de acesso à água
usobanheiro2	Variável dummy para domicílios que partilham o uso do banheiro ¹³
usobanheiro10	Variável dummy para falta de informação sobre uso do banheiro
telefonefixo4	Variável dummy para domicílios sem acesso ao telefone fixo ¹⁴
telefonefixo10	Variável dummy para domicílios sem informação sobre o acesso ao telefone fixo
raçachefe2	Variável dummy para raça do chefe (branco= 0 e negro = 1)
sexchefe4	Variável dummy para gênero do chefe (masculino = 0 e feminino = 1)
banhlixo2	Variável dummy para banheiro utilizado somente pelo domicílio e o destino do lixo inadequado
banhlixo3	Variável dummy para banheiro utilizado somente pelo domicílio e não existe informação sobre o destino do lixo
banhlixo4	Variável dummy para banheiro utilizado por diversos domicílios e destino do lixo adequado
banhlixo5	Variável dummy para banheiro utilizado diversos domicílios e destino do lixo inadequado
banhlixo6	Variável dummy para banheiro utilizado diversos domicílios e não existem informações sobre destino do lixo
banhlixo10	Variável dummy para domicílios sem informação sobre o uso do banheiro e o destino do lixo é adequado
banhlixo11	Variável dummy para domicílios sem informação sobre o uso do banheiro e o destino do lixo é inadequado
banhlixo12	Variável dummy para domicílios sem informação sobre o uso do banheiro e o destino do lixo
maquilavar4	Variável dummy para domicílio sem máquina de lavar roupa ¹⁵
maquilavar10	Variável dummy para domicílio sem informação sobre a posse de máquina de lavar roupa
matparede2	Variável dummy para domicílio cujo material da parede é pouco adequado ¹⁶
matcobertura2	Variável dummy para domicílio cujo material de cobertura não é durável
matcobertura3	Variável dummy para domicílio cujo material de cobertura é inadequado
matcobertura3	Variável dummy para domicílio em que falta informação sobre material de cobertura ¹⁷
região2	Variável dummy para região Nordeste ¹⁸
região3	Variável dummy para região Sudeste
região4	Variável dummy para região Sul
região5	Variável dummy para região Centro-Oeste
areacens2	Variável dummy para domicílios localizados no urbano interior (modelo logit 1) ou Variável dummy para domicílios localizados no rural interior (modelo logit 2) ¹⁹
destinoliço2	Variável dummy para domicílios cujo destino do lixo é inadequado ²⁰

O modelo também contempla a variável acesso ao telefone fixo (telefonefixo) e a variável relacionada à posse de máquina de lavar roupa (maquilavar). As variáveis de localização geográfica são expressas por meio da variável região (região) e da variável área censitária (areacens). Algumas variáveis de interação também foram introduzidas no modelo, como a interação entre região geográfica e área censitária (regarecen) e a interação (banhlixo) entre as variáveis que captam o uso do banheiro no domicílio e o destino do lixo.

6. Impacto do Programa Bolsa-Família sobre Frequência e Atraso Escolar: Análise e Interpretação dos Resultados

Inicialmente interpretamos os resultados dos modelos logits utilizados nos pareamentos pelo escore de propensão e, em seguida, analisamos os resultados dos testes de diferenças dos indicadores educacionais entre grupo de tratamento e controle.

6.1. Modelo logit

A partir do levantamento de diversas variáveis relacionadas à situação de pobreza das famílias, sobretudo aquelas vinculadas aos recursos familiares, realizou-se diversos testes para se obter o modelo mais adequado para o procedimento do pareamento pelo escore de propensão.

¹⁰ A categoria de referência corresponde a família com chefe e cônjuge.

¹¹ A categoria de referência corresponde à região Norte e área censitária interior.

¹² Essa variável é obtida a partir da variável V0212 da PNAD. A categoria de referência corresponde aos domicílios com acesso adequado à água. As condições inadequadas são definidas pela água proveniente de Poço, Nascente e Outra origem enquanto as condições são adequadas se água proveniente de rede geral de distribuição.

¹³ A categoria de referência corresponde ao domicílio cujo uso do banheiro se destina somente ao próprio domicílio.

¹⁴ A categoria de referência corresponde ao domicílio que tem acesso ao telefone fixo.

¹⁵ A categoria de referência corresponde ao domicílio que tem máquina de lavar roupa.

¹⁶ Essa variável foi construída a partir da variável V0203 do dicionário de domicílio da PNAD/2006. A categoria de referência corresponde ao domicílio cujo material da parede é adequado. O uso de zinco na construção das paredes é considerado pouco adequado, enquanto que o uso de Madeira aproveitada, de Palha e de Outro material é classificado como inadequado.

¹⁷ Essa variável foi construída a partir da variável V0204 do dicionário de domicílio da PNAD/2006. A categoria de referência corresponde ao domicílio cujo material da cobertura é adequado. O uso de zinco na construção das paredes é considerado pouco adequado, enquanto que a utilização de Madeira aproveitada, de Palha e de Outro material se considera inadequado. Por outro lado, se considera que o material de cobertura da moradia adequado desde que seja de telha, laje ou madeira aparelhada.

¹⁸ A categoria de referência corresponde a região Norte.

¹⁹ A categoria de referência corresponde a área censitária metropolitana.

²⁰ Essa variável é construída a partir da variável domiciliar V0218. A categoria de referência corresponde aos domicílios cujo destino do lixo é adequado. O lixo tem destino adequado quando sua coleta se verifica de forma direta ou indireta, enquanto o destino inadequado compreende lixo queimado ou enterrado na propriedade, jogado em terreno baldio ou logradouro, jogado em rio, lago ou mar ou outro destino.

De acordo com a Tabela 1, algumas variáveis eleitas não se revelaram significativas, mas segundo Rubin e Thomas (1996) uma variável somente deve ser eliminada do modelo caso se tenha convicção de sua insignificância para o resultado do modelo. Posto isso, optamos por manter todas as variáveis independentes na construção do modelo logit.

De início, interpreta-se os resultados do modelo logit utilizado na estimativa da probabilidade de participação no programa para as famílias residentes no meio urbano. As variáveis independentes têm o sinal esperado, porém vamos nos ater somente àquelas que se revelaram significativas em nível de 0,1%, 1% ou 5%. O coeficiente da variável rendimento demonstra que quanto menor o nível da renda familiar per capita maior a probabilidade de participar dos programas (Tabela 1). Verifica-se também uma relação inversa entre as chances de participação no programa e as variáveis que expressam anos de estudo da pessoa de referência, aquelas que se referem às famílias com presença somente do chefe e a outros arranjos familiares, as que tratam a ausência de informação sobre o uso do banheiro e aquelas que se referem aos residentes nas metrópoles do Sudeste ou aos moradores da região Centro-Oeste. Um resultado intrigante se refere à menor possibilidade de adesão ao programa nas famílias que têm somente o responsável. Esse resultado pode ser atribuído ao menor capital social existente nessas famílias e que pode inclusive dificultar sua organização e informação acerca das políticas sociais.

Por outro lado, as famílias sem telefone fixo ou sem máquina de lavar roupas têm maiores possibilidades de inclusão no Bolsa-Família. Além disso, famílias com chefes de cor negra, bem como aquelas residentes em habitações em que não existem informações sobre o uso do banheiro mas o destino dos dejetos se revela adequado, as famílias que moram no Nordeste, Sudeste ou Sul, as famílias que residem no interior e aquelas em cujo domicílio o destino do lixo se mostra inadequado apresentam também maiores chances de inclusão no programa.

O modelo logit utilizado no cômputo das chances das famílias moradoras na área rural participarem do programa aponta novamente que o maior nível de renda familiar ou de escolaridade do chefe da família reduz as probabilidades de participação no programa (Tabela 2). Também têm menores possibilidades de inclusão no programa as famílias classificadas como outros tipos de organização familiar, bem como as famílias residentes em domicílios cujo acesso à água se revela inadequado, naqueles em que não existem informações sobre o uso do banheiro e nos domicílios cujo material de cobertura se mostra inadequado. Essas famílias da zona rural têm déficit de capital social e de acesso a infraestrutura. Nesse caso, a gravidade da situação de precariedade pode impedir inclusive sua organização e informação acerca das políticas públicas. Ademais, as famílias chefiadas por pessoas do sexo feminino ou residentes na zona rural do Centro-Oeste têm menores chances de participarem do programa de transferência de renda.

As famílias residentes na área rural têm maior probabilidade de participar do Bolsa-Família se não possuem máquina de lavar roupa, telefone fixo no domicílio ou informação sobre a posse desse meio de comunicação, bem como se a pessoa de referência for negra. Além disso, os domicílios sem informação sobre o uso

Tabela 3

Modelo logit para o cálculo da probabilidade de participação no programa Bolsa-Família para famílias residentes no meio urbano*

Variáveis	Coefficientes	$p > z $
yfamllqpc	-0.0055	0.000
Itipofam 2	-0.4992	0.000
Itipofam 3	-0.7353	0.000
Iregarecens sum 3	-0.1481	0.159
Iregarecens sum 5	-0.4485	0.000
Iregarecens sum 8	-0.1579	0.238
Iregarecens sum 9	0.0784	0.617
Anoestupai sum	-0.0658	0.000
Iorigemagua sum 2	-0.0211	0.753
Iorigemagua sum 10	0.0295	0.658
Iusobanheiro sum 2	0.1848	0.147
Iusobanheiro sum 10	-154.508	0.000
Itelefonefixo sum 4	0.4416	0.000
Itelefonefixo sum 10	-160.359	.
Iraçachefe sum 2	0.1968	0.000
Isexchefe sum 4	0.0900	0.124
Ibanhlixo sum 2	-151.352	0.000
Ibanhlixo sum 5	-140.551	0.000
Ibanhlixo sum 10	155.212	0.000
Imaquinalavar sum 4	0.4914	0.000
Imaquinalavar sum 10	477.669	.
Imatparede sum 2	0.0742	0.514
Imatcobertura sum 2	0.2225	0.134
Imatcobertura sum 3	-0.1775	0.493
Imatcobertura sum 10	-160.999	.
Iregião 2	0.5438	0.000
Iregião 3	0.2751	0.000
Iregião 4	0.2415	0.032
Iregião 5	-0.2309	0.002
Iareacens 2	0.2498	0.008
Idestinolixo sum 2	1.540.741	0.000
cons	-0.5835	0.000

*Variável dependente é dicotômica com código 1 para famílias que participam dos programa e 0 em caso contrário. (Elaboração própria).

do banheiro e destino inadequado do lixo ou moradores do interior do Nordeste também têm maiores chances de inclusão no programa.

Cabe observar que no modelo logit para os residentes na zona urbana, os testes de diferenças médias entre as covariáveis antes e após o tratamento não se revelaram significativos em nível 0,1%, 1% e 5%. Ou seja, o modelo satisfaz a hipótese de balanceamento. Se comparamos o pseudo- R^2 para amostra antes e após o

Tabela 4

Modelo logit para o cálculo da probabilidade de participação no programa Bolsa-Família para famílias residentes no meio rural*

Variáveis	Coefficientes	$p > z $
Yfam1qpc	-0.0076	0.000
Itipofam 2	-0.4093	0.056
Itipofam 3	-11.833	0.000
Iregarecens sum 2	0.2350	0.652
Iregarecens sum 3	0.3403	0.420
Iregarecens sum 4	0.5395	0.000
Iregarecens sum 6	0.4889	0.281
Iregarecens sum 8	-0.3877	0.362
Anoestupai sum	-0.0394	0.000
Iorigemagua sum 2	-0.2083	0.023
Iorigemagua sum 10	-0.2504	0.010
Iusobanheiro sum 2	-11.687	0.104
Iusobanheiro sum 10	-171.670	0.000
Itelefonefixo sum 4	0.8478	0.000
Itelefonefixo sum 10	174.568	0.000
Iraçachefe sum 2	0.1821	0.009
Isexchefe sum 4	-0.5239	0.007
Ibanhlixo sum 2	-13.155	0.128
Ibanhlixo sum 10	172.694	.
Ibanhlixo sum 11	160.321	0.000
Imaquinalavar sum 4	0.3689	0.011
Imatparede sum 2	0.1793	0.159
Imatcobertura sum 2	0.1559	0.534
Imatcobertura sum 3	-0.5958	0.000
Iregião 3	-0.0771	0.866
Iregião 4	0.2173	0.612
Iregião 5	-0.5247	0.000
Iareacens 2	0.1808	0.620
Idestinolixo sum 2	15.152	0.620
cons	-0.4296	0.326

*Variável dependente é dicotômica com código 1 para famílias que participam dos programa e 0 em caso contrário. (Elaboração própria).

pareamento pelo método do vizinho mais próximo com reposição, seu valor passa de 0.113 para 0.001. Segundo Sianesi (2004), essa redução seria aguardada porque a maior homogeneidade entre a amostra reduz o poder explicativo do modelo que resulta de um pareamento de boa qualidade.

No caso do modelo logit para os residentes na zona rural, os testes de diferenças de médias para todas as variáveis independentes mais uma vez não se revelou significativo em nível de 0,1%, 1% e 5%. Se comparamos o pseudo- R^2 para amostra antes e após o pareamento pelo método do vizinho mais próximo com reposição,

seu valor passa de 0.146 para 0.004. Em resumo, os modelos logit utilizados tiveram desempenhos satisfatórios no testes de balanceamento.

6.2. Testes de Diferenças das Médias dos Indicadores Educacionais

Após os pareamentos para famílias residentes nas áreas urbana e rural, realizou-se o confronto das médias dos indicadores educacionais. Vale observar que a realização de testes de diferenças de médias que considerem o desenho da PNAD como amostra aleatória simples tem estimativas de erros padrões com viés para baixo, assim optou-se pelo uso de comandos do *Stata* que considerem o desenho da amostra dessa pesquisa como estratificado. Posto isto, o cálculo da média dos indicadores educacionais para amostra estratificada foi obtido via o comando do *Stata* *svy linearized: mean*. Os testes de significância das diferenças de médias dos indicadores educacionais foram obtidos por meio do comando do *Stata* *lincom*.

Nesta seção somente apresentaremos os resultados obtidos para os grupos de tratamento e controle definidos por meio do algoritmo do vizinho mais próximo. Encontram-se no Apêndice os resultados dos testes de médias realizados para as amostras obtidas por meio dos demais algoritmos (Tabelas 5 a 8), bem como os resultados dos testes estimados via bootstrapping (Tabelas 9 e 10).

No Brasil, as estatísticas descritivas apontaram que na década atual a quase totalidade das crianças e adolescentes entre 7 e 15 anos de idade, de distintos espaços geográficos e faixas de rendimento, frequentavam a escola, porém outros indicadores como o atraso escolar e a evasão permanecem elevados.

Posto isso, não se aguardava que o Programa Bolsa-Família elevasse a frequência escolar, entendida aqui como a taxa de atendimento, porém algum impacto poderia ser esperado na redução da ausência em sala de aula e na defasagem idade-série. O programa reduziria o número de faltas no ano letivo em função de sua condicionalidade, que exige a frequência em 85% das aulas. Em relação ao indicador defasagem idade-série, a contribuição do programa poderia se verificar por meio de alguns canais de transmissão. Um deles se refere à redução do custo indireto da educação por meio da ajuda financeira às famílias, que ao estimular a permanência na escola poderia combater um dos principais responsáveis pelo atraso escolar, qual seja, a evasão. Outra contribuição do programa adviria do estímulo que a ajuda financeira criaria para a entrada das crianças na escola na idade prevista pela legislação educacional. Além do que, a literatura aponta que crianças e adolescentes mais assíduos às aulas têm melhor desempenho escolar e, portanto, realizam a progressão escolar e evitam o atraso escolar (Barbosa 2009). Todavia, a defasagem idade-série tem múltiplos determinantes, como os fatores individuais observáveis e não observáveis, as características socioeconômicas das famílias e a própria escola. Neste artigo, nos concentramos na contribuição dos recursos familiares, em especial da renda familiar, para o atraso escolar.

6.2.1. Ausência escolar

Como se vê na Tabela 5, os testes de diferenças de médias das razões entre as crianças e adolescentes residentes nas áreas urbana e rural, dos grupos de tratamento e controle não foram significativos.²¹ Os estudantes que integram o grupo de tratamento e o de controle desses recortes geográficos tiveram ausência escolar semelhante, no período de 2 de agosto a 30 de setembro de 2006.

Tabela 5

Resultados dos Testes de diferenças de médias das razões de crianças e adolescentes residentes na zona urbana e rural que deixaram de comparecer pelo menos 1 dia à escola no período de 2 de agosto a 30 de setembro de 2006 nas famílias do grupo de controle e de tratamento

Localização	Diferenças das médias*	Erro padrão	<i>t</i>	<i>P</i> > <i>t</i>	Intervalo de confiança de 95%	
Residentes na zona urbana	0.0069799	0.0113328	0.62	0.538	-0.0152419	0.0292017
Residentes na zona rural	0.0167058	0.0240245	0.70	0.487	-0.0305198	0.0639315

*Resultado das diferenças entre as médias das razões de crianças e adolescentes que se ausentaram das aulas no grupo de controle e tratamento. (Elaboração própria).

Nota-se que no caso dessa variável o programa não gerou uma distinção entre os beneficiários e não beneficiários, portanto o Bolsa-Família se revela inócuo. Os demais algoritmos utilizados nesse estudo também confirmaram que as divergências entre o grupo de tratamento e o de controle, no que se refere à ausência escolar, não são significativas no espaço urbano nem no rural (Tabelas 7, 9 e 11 do Apêndice). Esse resultado confirma aquele encontrado por Romero e Hermeto (2009), em que a ausência escolar não teve divergência significativa entre o grupo de tratamento e controle em todos os recortes geográficos estudados.

6.2.2. Atraso escolar

Os testes de diferenças de médias das razões de crianças e adolescentes com defasagem idade-série entre o grupo de tratamento e o de controle revelaram que as divergências não são significativas tanto no meio urbano quanto no meio rural (Tabela 6).

Os pareamentos realizados pelo algoritmo do vizinho mais próximo com caliper ou pelo procedimento da distância de Mahalanobis, bem como as replicações realizadas via bootstrapping também confirmaram que não existem diferenças significativas nas razões de pessoas de 7 a 15 anos de idade com atraso escolar, no grupo de tratamento ou de controle, para famílias residentes na área urbana ou rural (Apêndice – Tabelas 8, 10 e 12).

O pressuposto aqui adotado assumiu que o programa Bolsa-Família contribuiria para elevar a assiduidade e reduzir o atraso escolar, na medida em que exigia um comparecimento de 85% às aulas por parte dos beneficiários na faixa etária

²¹ Não realizamos o teste de análise de sensibilidade que somente deve ser estimado em caso de resultados significativos.

Tabela 6

Teste de diferença das médias das razões de crianças e adolescentes residentes na zona urbana e rural com defasagem idade-série nas famílias do grupo de controle e de tratamento

Localização	Diferenças das médias*	Erro padrão	<i>t</i>	$P > t $	Intervalo de confiança de 95%	
Residentes na zona urbana	-0.0152328	0.0105931	-1.44	0.151	-0.0360043	0.0055388
Residentes na zona rural	0.0206353	0.0234485	0.88	0.379	-0.025458	0.0667285

*Resultado das diferenças entre as médias das razões de crianças e adolescentes com atraso escolar no grupo de controle e tratamento (Elaboração própria).

de 7 a 15 anos. Segundo os defensores da teoria do capital humano, o abandono escolar se deve à pobreza da família, que se expressa pelo nível de renda reduzido diante das necessidades básicas. Nessas famílias, qualquer membro que possa contribuir para a melhora do nível de rendimento é compelido ao trabalho, inclusive crianças e adolescentes. Nessa perspectiva, a transferência de assistência financeira condicionada à assiduidade das crianças pobres na escola promoveria uma redução no abandono escolar e, portanto, combateria um dos principais fatores responsáveis pelo atraso escolar.

Contudo, em 2008 somente 3% das pessoas entre 7 e 15 anos de idade declararam que não estudavam, além do que somente 19% dessas crianças e adolescentes trabalhavam. Ou seja, uma pequena parcela dessas pessoas que não estudava também trabalhava, assim as estatísticas descritivas sugerem que o trabalho infanto-juvenil não se revela o principal fator explicativo do abandono escolar. No ano de 2008, outro dado importante aponta que 33% das crianças e adolescentes tinham atraso escolar, sendo que 11% delas trabalhavam. Grande parte das crianças e jovens com defasagem idade-série, portanto, não trabalham. O programa Bolsa-Família se concentra em uma assistência financeira associada a uma contrapartida que é na verdade a assiduidade em sala de aula, porém os dados anteriores apontam que os resultados educacionais insatisfatórios existentes entre muitas crianças e jovens pobres e sintetizados na variável atraso escolar não podem ser atribuídos, sobretudo, ao trabalho infanto-juvenil que seria combatido indiretamente pelo programa.

Resta refletir sobre os motivos que possivelmente impediram que os canais de transmissão do programa Bolsa-Família atuassem sobre a defasagem idade-série. No caso da entrada da criança na idade preconizada pela legislação educacional, a repercussão do programa, provavelmente, seria reduzida, já que os dados apontam que, no ano de 2008, em torno de 98% das crianças com 7 anos de idade frequentavam a antiga primeira série ensino fundamental.

Leon e Menezes-Filho (2002) apontaram que existe um entrelaçamento entre evasão e repetência escolar. Ou seja, crianças que enfrentam a repetência são mais propensas ao abandono escolar. A ajuda monetária pode ter seus limites na reversão, sobretudo, da repetência, porque outros recursos familiares igualmente importantes não são contemplados pelo programa Bolsa-Família. Em geral, os

alunos mais assíduos têm melhor desempenho escolar, no entanto esse resultado pode também estar associado a outros recursos familiares, como o capital cultural da família. Para Boudon (1981), na etapa inicial da vida escolar a herança cultural das famílias influencia o desenvolvimento das habilidades básicas, porém nas etapas superiores aqueles que superaram os desafios reúnem as habilidades necessárias para persistir na trajetória escolar. Ou seja, em algum período da vida escolar, crianças inseridas em famílias com menor capital cultural enfrentam maiores déficits educacionais.

O programa Bolsa-Família almeja solucionar a pobreza no longo prazo com a melhora nos indicadores educacionais que, por sua vez, tem vários determinantes, desde características individuais à qualidade da escola. No âmbito de nosso estudo, vê-se que os recursos monetários são relevantes, já que os melhores indicadores escolares estão entre as crianças de famílias mais ricas, porém o programa deve estar no mínimo articulado com outros programas que busquem compensar as famílias de seus vários déficits de recursos familiares, como o déficit em capital cultural.

7. Notas Conclusivas

De acordo com as PNADs, a média de estudo da população brasileira com 15 anos ou mais saltou de 5,2 anos em 1992 para 7,4 anos em 2008. Após mais de uma década elevou-se em dois anos essa média que ainda permanece baixa. Segundo a perspectiva dos teóricos do capital humano, a origem desse quadro educacional se deve aos rendimentos que as famílias pobres podem obter com o trabalho infanto-juvenil. Essas crianças e jovens entram precocemente no mercado de trabalho e, provavelmente, abandonam a escola, e o resultado final é um baixo nível de escolaridade. Na linha de raciocínio dos defensores da teoria do capital humano, os anos de estudo determinam a produtividade do trabalho, que por sua vez determina os rendimentos recebidos. Um caminho para o rompimento desse círculo vicioso da pobreza seria a transferência monetária, no intuito de compensar as famílias da renda proveniente do trabalho infantil. Nesse raciocínio, a principal razão para as crianças e jovens não frequentarem a escola se deve ao nível de renda das famílias ou à pobreza absoluta de seus familiares. A elevação desse nível de renda por meio de políticas sociais focalizadas seria condição necessária e suficiente para a superação dos baixos índices de escolaridade. Em resumo, esses teóricos se concentram no nível de rendimento familiar para explicar o baixo nível de escolaridade existente entre muitos brasileiros.

A vertente da teoria do capital humano embasa teoricamente o programa Bolsa-Família. Ou seja, a transferência monetária às famílias, associada à condição das crianças e jovens frequentarem as aulas, contribuiria para elevar o nível de escolaridade dessas pessoas que, uma vez adultas, obteriam maiores rendimentos.

Os resultados dos nossos testes de diferenças entre as proporções médias de crianças que se ausentaram das aulas, nos grupos de tratamento e de controle,

não foram significativos tanto no meio urbano quanto no meio rural. No caso do indicador de defasagem idade-série, os testes de diferenças entre as razões médias de crianças e adolescentes com atraso escolar, no grupo de tratamento e no grupo controle, não foram também significativos em qualquer dos recortes espaciais.

As estatísticas descritivas mencionadas acima revelaram que uma pequena proporção das crianças e jovens com atraso escolar trabalhavam, assim não seria provavelmente o trabalho o motivo central e exclusivo da performance escolar insatisfatória entre muitos alunos pobres no país. Conforme salienta Bourdieu (1999a), o êxito e a trajetória escolar se vincula ao capital cultural das famílias, entendido como anos de estudo, facilidade verbal, acesso aos bens culturais como livros, visitas a museus etc. Noutras palavras, uma concepção ampla de recursos familiares, em especial do capital cultural, pode estabelecer nexos causais importantes entre esses recursos e o desempenho escolar.

Para o autor, a herança cultural das famílias se revela decisiva na compreensão de suas atitudes em face do sistema educacional. As crianças e jovens que vivem em famílias com elevado capital cultural têm familiares mais bem informados sobre as possibilidades e exigências escolares e contam com a possibilidade de ajuda por parte de seus pais nas tarefas escolares. Enfim, as pessoas inseridas nessas famílias possuem informações, acessos e ajudas familiares que são decisivas em sua trajetória escolar. Na perspectiva de Bourdieu (1999a,b,c), o fracasso escolar de muitas crianças e adolescentes se deve ao déficit de capital cultural existente em diversas famílias, que deveria ser compensado pelas instituições escolares. Desse modo, a permanência na escola, a realização das transições escolares, a redução da repetência e a eliminação da defasagem idade-série exigem um conjunto de ações que não podem ser reduzidas à mera transferência de renda para as famílias pobres, na medida em que o recurso monetário não compensaria os outros déficits de recursos familiares igualmente importantes na vida escolar das crianças. Boundon (1981) concordaria com os argumentos de Bourdieu, no entanto restringiria sua validade para as etapas iniciais da vida escolar. À parte a polêmica, existe um consenso sobre a importância da herança cultural para o êxito da trajetória escolar das crianças, ainda que somente na etapa inicial. Ou seja, é importante que existam mecanismos de compensação desse déficit na vida escolar das crianças pobres.

Os resultados deste estudo revelam para os indicadores educacionais eleitos que o programa Bolsa-Família se mostrou inócuo. Os testes não confirmaram nosso pressuposto de uma possível melhora nesses indicadores, já que o programa exigia alto percentual de frequência às aulas. Em geral, os alunos mais assíduos apresentam melhor performance educacional, porém existem fatores correlacionados com a maior frequência às aulas, como a atitude da família em relação à escola, o capital cultural dessas famílias, que também devem contribuir para os bons resultados alcançados por essas crianças e adolescentes.

Cabe observar, no entanto, que o programa pode contribuir para elevar os anos de estudo entre as crianças e jovens pobres, na medida em que estimule a sua permanência por um período maior na escola, porém esse resultado não será decorrente de uma melhora nos indicadores de performance educacional. Em

resumo, o programa não se encontra articulado a nenhum outro que busque superar os diversos déficits de recursos entre as famílias pobres. É necessária, portanto, uma reformulação do programa de modo a inseri-lo em um projeto que busque elevar o nível de escolaridade de crianças e jovens por meio da melhora de sua performance escolar e que, portanto, os ajude na superação de seus déficits. A ajuda financeira concedida às famílias pobres se revela uma condição necessária, mas insuficiente para a obtenção da melhora dos indicadores educacionais entre as crianças e jovens pobres e, portanto, para a superação da pobreza no longo prazo.

Referências bibliográficas

- Barbosa, M. L. O. (2009). *Desigualdade e Desempenho: Uma Introdução à Sociologia da Escola Brasileira*. Argvmentvm, Belo Horizonte.
- Berhrman, J., Sengupta, P., & Todd, P. (2001). Progressing through Progres: An impact assesment of a school subsidy experiment in Mexico. PIER working paper 01-033, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania. Disponível em: <http://www.econ.upenn.edu/system/files/01-033.pdf>. Acesso em: 25 mar. 2006.
- Boudon, R. (1981). *A Desigualdade das Oportunidades: A Mobilidade Social nas Sociedades Industriais*. Editora Universidade de Brasília, Brasília.
- Bourdieu, P. (1999a). A escola conservadora: As desigualdades frente à escola. In Nogueira, M. A. & Catani, A., editors, *Escritos de Educação*, pages 39–64. Vozes, Petrópolis.
- Bourdieu, P. (1999b). O capital social – Notas introdutórias. In Nogueira, M. A. & Catani, A., editors, *Escritos de Educação*, pages 65–70. Vozes, Petrópolis.
- Bourdieu, P. (1999c). Os três estados do capital cultura. In Nogueira, M. A. & Catani, A., editors, *Escritos de Educação*, pages 71–80. Vozes, Petrópolis.
- Cacciamali, M. C., Tatei, F., & Batista, N. N. F. (2009). Pobreza, trabalho infantil e programa bolsa família. In Cacciamali, M. C. & José-Silva, M. F., editors, *A Construção da Igualdade de Gênero e Raça na América Latina no Século XXI: O Caso Brasil*, pages 81–113. Suprema, São Carlos.
- Caliendo, M. & Kopeinig, S. (2005). Some practical guidance for the implementation of propensity score matching. IZA Discussion 1588, Institute for the Study of Labor.
- Davidson, R. & Mackinnon, J. G. (2000). Bootstrap tests. How many bootstraps? *Econometric Reviews*, 19:55–68.
- Duryea, S. & Morrison, A. (2004). The effect of conditional transfers on school performance and child labor: Evidence from an ex-post impact evaluation in Costa Rica. In *Inter-American Development Bank*. Disponível em: <http://www.iadb.org/res/publications/pubfiles/pubWP-505.pdf>. Acesso em: 23 set. 2006.
- Ferro, A. R. & Kassouf, A. L. (2005). Avaliação do impacto dos programas bolsa-escola sobre o trabalho infantil no Brasil. *Pesquisa e Planejamento*, 35:417–444.
- Heckman, J., Ichimura, J. S., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, 66:1017–1098.
- Leon, F. L. L. & Menezes-Filho, N. A. (2002). Reprovação, avanço e evasão escolar no Brasil. *Pesquisa e Planejamento Econômico*, 32:417–451.
- Machado, D. C. & Gonzaga, G. (2007). O impacto dos fatores familiares sobre a defasagem idade-série de crianças no Brasil. *Revista Brasileira de Economia*, 61:449–476.

- Neder, H. D. (2008). *Amostragem em Pesquisas Socioeconômicas*. Alínea, Campinas.
- Resende, A. C. C. (2006). Avaliando resultados de um programa de transferência de renda: O impacto do bolsa-família sobre os gastos das famílias brasileiras. Master's thesis, CEDEPLAR, Universidade Federal de Minas Gerais, Belo Horizonte.
- Romero, J. A. R. & Hermeto, A. M. (2009). Avaliação de impacto do programa bolsa família sobre indicadores educacionais: Uma abordagem de regressão descontínua. In ANPEC, editor, *Anais do Encontro Nacional de Economia*, Foz do Iguaçu.
- Rosenbaum, P. & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–50.
- Roy, A. (1951). Some thoughts of the distribution of earning. *Oxford Economic Papers*, 3:135–145.
- Rubin, D. (1974). Estimating causal effects to treatments in randomised and nonrandomised studies. *Journal of Educational Psychology*, 66:688–701.
- Rubin, D. B. & Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics*, 52:249–264.
- Sianesi, B. (2004). An evaluation of the active labour market programmes in Sweden. *The Review of Economics and Statics*, 86:133–155.
- Silva, N. V. (2003). Expansão escolar e estratificação educacional no Brasil. In Hasenbalg, C. & Silva, N. V., editors, *Origens e Destinos: Desigualdades Sociais ao Longo da Vida*, pages 105–146. Topbooks.
- Soares, F. V., Soares, S., & Medeiros, M. (2006). Programa de transferência de renda no Brasil: Impactos sobre a desigualdade. Texto para Discussão 1228, IPEA, Brasília.
- Soares, S., Osório, R. G., Soares, F. V., Medeiros, M., & Zepeda, E. (2007). Programa de transferência de renda no Brasil, Chile e México: Impactos sobre a desigualdade. Texto para Discussão 1293, IPEA, Brasília.
- Weiss, A. (1988). High school graduation, performance and wages. *Journal of Political Economy*, 96:785–820.

Apêndice

A1. Pareamento via algoritmo do vizinho mais próximo com caliper

Tabela 7

Resultados dos Testes de diferenças de médias das razões de crianças e adolescentes residentes na zona urbana e rural que deixaram de comparecer pelo menos 1 dia à escola no período de 2 de agosto a 30 de setembro de 2006 nas famílias do grupo de controle e tratamento

Localização	Diferenças das médias*	Erro padrão	t	$P > t $	Intervalo de confiança de 95%	
Residentes na zona urbana	.0069799	.011346	0.62	0.538	-.0152678	.0292276
Residentes na zona rural	.0167058	.0240245	0.70	0.487	-.0305198	.0639315

*Resultado das diferenças entre as médias das razões de crianças e adolescentes que se ausentaram das aulas no grupo de controle e tratamento.

Tabela 8

Teste de diferença das médias das razões de crianças e adolescentes residentes na zona urbana e rural com defasagem idade-série nas famílias do grupo de controle e tratamento

Localização	Diferenças das médias*	Erro padrão	t	$P > t $	Intervalo de confiança de 95%	
Residentes na zona urbana	-.0152328	.0105409	-1.45	0.149	.0359021	.0054365
Residentes na zona rural	.0206353	.0234479	0.88	0.379	-.0254568	.0667273

*Resultado das diferenças entre as médias das razões de crianças e adolescentes com atraso escolar no grupo de controle e tratamento.

A2. Pareamento via distância de Mahalanobis

Tabela 9

Resultados dos Testes de diferenças de médias das razões de crianças e adolescentes residentes na zona urbana e rural que deixaram de comparecer pelo menos 1 dia à escola no período de 2 de agosto a 30 de setembro de 2006 nas famílias do grupo de controle e tratamento

Localização	Diferenças das médias*	Erro padrão	t	$P > t $	Intervalo de confiança de 95%	
Residentes na zona urbana	-.0036271	.0120026	-0.30	0.763	-.0271623	.0199081
Residentes na zona rural	.0254832	.0209486	1.22	0.225	-.0156964	.0666627

*Resultado das diferenças entre as médias das razões de crianças e adolescentes que se ausentaram das aulas no grupo de controle e tratamento.

Tabela 10

Teste de diferença das médias das razões de crianças e adolescentes residentes na zona urbana e rural com defasagem idade-série nas famílias do grupo de controle e tratamento

Localização	Diferenças das médias*	Erro padrão	t	$P > t $	Intervalo de confiança de 95%	
Residentes na zona urbana	-.016432	.0111383	-1.48	0.140	-.0382727	.0054087
Residentes na zona rural	-.0154692	.0220345	-0.70	0.483	-.0587827	.0278443

*Resultado das diferenças entre as médias das razões de crianças e adolescentes com atraso escolar no grupo de controle e tratamento.

A3. Bootstrapping

Tabela 11

Resultados dos Testes de diferenças de médias das razões de crianças e adolescentes residentes na zona urbana e rural que deixaram de comparecer pelo menos 1 dia à escola no período de 2 de agosto a 30 de setembro de 2006 nas famílias do grupo de controle e tratamento

Localização	Diferenças das médias*	Erro padrão	t	$P > t $	Intervalo de confiança de 95%	
Residentes na zona urbana	0.0069799	0.0093087	0.75	0.453	-0.0112648	0.0252246
Residentes na zona rural	0.0167058	0.0181783	0.92	0.358	-0.0189231	0.0523347

*Resultado das diferenças entre as médias das razões de crianças e adolescentes que se ausentaram das aulas no grupo de controle e tratamento.

Tabela 12

Teste de diferença das médias das razões de crianças e adolescentes residentes na zona urbana e rural com defasagem idade-série nas famílias do grupo de controle e tratamento

Localização	Diferenças das médias*	Erro padrão	t	$P > t $	Intervalo de confiança de 95%	
Residentes na zona urbana	-0.0152328	0.0083871	-1.82	0.069	0.0316713	0.0012057
Residentes na zona rural	0.0206353	0.0175722	1.17	0.240	-0.0138057	0.0550762

*Resultado das diferenças entre as médias das razões de crianças e adolescentes com atraso escolar no grupo de controle e tratamento.