

Universidade de São Paulo
Faculdade de Economia, Administração e Contabilidade de Ribeirão
Preto
Departamento de Economia
Programa de Pós-Graduação em Economia

Pedro Cavalcanti de Camargo

Uma análise do efeito do *Programa Bolsa Família* sobre o
desempenho médio das escolas brasileiras

Orientador: Profa. Dra. Elaine Toldo Pazello

Ribeirão Preto

2011

Prof. Dr. João Grandino Rodas
Reitor na Universidade de São Paulo

Prof. Dr. Sigismundo Bialoskorski Neto
Diretor de Faculdade de Economia, Administração e Contabilidade de
Ribeirão Preto

Prof. Dr. Walter Belluzzo Jr
Chefe do Departamento de Economia

Prof. Dr. Alex Luiz Ferreira
Coordenador do Programa de Pós-Graduação em Economia

Pedro Cavalcanti de Camargo

Uma análise do efeito do *Programa Bolsa Família* sobre o
desempenho médio das escolas brasileiras

Dissertação Apresentada ao Programa de Pós-Graduação em Economia – Área: Economia Aplicada da Faculdade de Economia, Contabilidade e Administração de Ribeirão Preto como requisito para obtenção do título de Mestre em Ciências.

Orientador: Profa. Dra. Elaine Toldo Pazello

Ribeirão Preto

2011

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada à fonte.

Camargo, Pedro Cavalcanti de

Uma análise do efeito do *Programa Bolsa Família* sobre o desempenho médio das escolas brasileiras. Ribeirão Preto, 2011.

58 f. il.; 30 cm

Dissertação de Mestrado, apresentada à Faculdade de Economia, Contabilidade e Administração de Ribeirão Preto/USP. Área de concentração: Economia Aplicada.

Orientador: Pazello, Elaine Toldo

1. Políticas Públicas 2. Programa Bolsa Família. 3. Economia da Educação. 4. Desempenho Escolar.

Folha de Aprovação

Nome: CAMARGO, Pedro Cavalcanti de
Título: Uma análise do efeito do
Programa Bolsa Família sobre o
desempenho médio das escolas
brasileiras

Dissertação Apresentada ao Programa de Pós-
Graduação em Economia – Área: Economia
Aplicada da Faculdade de Economia,
Contabilidade e Administração de Ribeirão
Preto como requisito para obtenção do título de
Mestre em Ciências.

Aprovado em: _____

Banca Examinadora

Prof.: _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof.: _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof.: _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Agradecimentos

Primeiramente, agradeço à minha família, que me apoiou incondicionalmente ao longo de toda minha vida. Sem a presença de vocês jamais conseguiria chegar onde estou.

Aos meus queridos amigos que, ao longo destes anos todos, me proporcionaram muita alegria. Aproveito também para prestar uma homenagem a um amigo muito especial que nos deixou há alguns anos.

Aos professores do PPGE da FEA-RP, especialmente a minha orientadora Profa. Dra. Elaine Pazello pelo conhecimento que pude adquirir neste importante estágio da minha vida acadêmica.

A CAPES – Observatório da Educação - pelo apoio financeiro para a realização deste trabalho.

Por fim, agradeço a Andreia que esteve ao meu lado durante um dos maiores passos que dei na minha vida.

Resumo

CAMARGO, P. C. ***Uma análise do efeito do Programa Bolsa Família sobre o desempenho médio das escolas brasileiras.*** 2011. 58 f. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto. Universidade de São Paulo.

Este trabalho objetiva avaliar o efeito de um aumento da proporção de alunos beneficiados pelo *Programa Bolsa Família* nas escolas, sobre o desempenho médio das mesmas. Utilizando dados do Censo Escolar 2008, da “Frequência Escolar dos beneficiários do *Programa Bolsa Família*” de 2008 e da Prova Brasil 2009 este texto procura relacionar o desempenho médio por escola dos alunos de 4ª série com a proporção de alunos participantes do programa. Para avaliar o desempenho das escolas, este trabalho utilizou as notas médias nos exames de Português e Matemática da Prova Brasil, a taxa de aprovação e a taxa de abandono escolar. Os resultados obtidos sugerem que escolas com maior proporção de alunos tratados, em média, possuem menores taxas de abandono e de aprovação. Para os exames de proficiência, os resultados também apontam para uma relação negativa, entretanto há evidências de que escolas que viriam a receber uma quantidade maior de alunos beneficiados pelo PBF já possuíam notas médias nestes exames inferiores àquelas com proporção menor antes da criação do programa. Desta forma, este trabalho apresenta evidências de que há uma necessidade de maior atenção por parte dos gestores de políticas públicas com as escolas com maior porcentagem de alunos beneficiados pelo *Programa Bolsa Família*.

Palavras-Chave: Educação, *Programa Bolsa Família*, desempenho escolar.

Abstract

CAMARGO, P. C. *An analysis of the effect of the Programa Bolsa Família on the average performance of schools in Brazil*. 2011. 58 f. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto. Universidade de São Paulo.

This text aims to evaluate the effect of an increase in the percentage of students treated by the *Programa Bolsa Família* in each school on the average performance of that same school. Using data from Censo Escolar 2008, “Frequência Escolar dos beneficiários do *Programa Bolsa Família*” for 2008 and Prova Brasil 2009 this paper seeks to relate the average performance of each school with the proportion of students participating on the program. To evaluate the performance of schools, this study used the average scores on Portuguese and Mathematics in Prova Brasil’s tests, the approval rate and dropout rate for fourth graders. The result suggests that schools with higher proportion of treated students, on average, have lower dropout rates and lower approval rates. The results also point to a negative relationship between the proportion of benefited pupils and proficiency tests; however there is evidence that schools who would receive a larger amount of treated students had inferior average scores on these tests before the creation of the program. Thus, this paper presents evidence that there is a need for greater attention from policy makers with schools with higher percentage of students benefited from the *Programa Bolsa Família*.

Key-words: Education, *Programa Bolsa Família*, school performance.

Sumário

1. INTRODUÇÃO	13
2. DESCRIÇÃO DAS BASES DE DADOS UTILIZADAS E ANÁLISE DESCRITIVA DOS DADOS OBTIDOS	21
3. METODOLOGIA.....	30
4. RESULTADOS	35
4.1. ESTIMAÇÃO DA PROBABILIDADE DO ALUNO RECEBER O TRATAMENTO	35
4.2. ESTIMAÇÃO DO EFEITO MARGINAL DE UM AUMENTO DA PROPORÇÃO DE ALUNOS PARTICIPANTES DO PBF	39
5. TESTE DE ROBUSTEZ DOS RESULTADOS ESTIMADOS	43
5.1. USO DE UMA BASE “PLACEBO”	43
5.2. DIFERENÇAS EM DIFERENÇAS	48
6. CONCLUSÃO.....	54
REFERÊNCIAS.....	56

Índice De Gráficos e Tabelas

Tabela 1: Proporção de alunos que recebem benefícios do PBF	23
Tabela 2: Desempenho dos alunos por escola em 2009	24
Tabela 3: Correlação entre o desempenho escolar e a proporção de beneficiados..	24
Tabela 4: Composição dos alunos nas escolas	25
Tabela 5: Características das Escolas por quartil de Propoção de beneficiados	26
Tabela 6: Proporção de escolas em cada quartil por região	28
Tabela 7: Estimação da probabilidade do aluno participar do PBF	35
Tabela 8: Estimativa de recebimento x Recebimento efetivo	36
Gráfico 1: Probabilidade de receber o benefício para tratados e não tratados	37
Tabela 9: Correlação entre a proporção de tratados por escola e a proporção estimada	38
Gráfico 2: Proporção de aluno tratados por escola x proporção estimada.....	38
Tabela 10: Estimativas de OLS para os resultados de interesse	39
Tabela 11: Estimativas de OLS para os resultados de interesse para diferentes modelos	40
Tabela 12: Estimativas de OLS utilizando o SAEB 2001 e Censo Escolar 2001	45
Tabela 13: Comparação entre os coeficientes do placebo e da base original	46
Tabela 14: Estimativa de Dif-em-dif por POLS.....	50
Tabela 15: Estimativa de Dif-em-dif utilizando novo D	51

1. Introdução

Desde o final da década de 1990, e principalmente durante o início dos anos 2000, muitos países em desenvolvimento, em especial os latino-americanos, adotaram políticas de transferência condicional de renda. A adoção de tais políticas de assistência à população nestes países se deve primordialmente aos elevados índices de pobreza e a alta desigualdade social, de forma que é natural que os governos federais destes países adotem políticas para reduzir tais índices. Embora se adequem às peculiaridades dos países para os quais foram projetados, estes programas contemplam alguns objetivos básicos comuns entre eles:

1. Dispor assistência financeira a famílias em situação de pobreza no curto prazo, de forma a reduzir a pobreza corrente e atenuar as desigualdades;
2. Incentivar o desenvolvimento de capital humano através das condicionalidades do recebimento do benefício, de maneira a diminuir a pobreza no longo prazo.

Dentre os mais conhecidos, pode-se citar o programa *Oportunidades* iniciado em 2002, baseado no seu predecessor *Progresá* de 1997 no México; o programa *Famílias em Acción* iniciado em 2001 na Colômbia, e o Programa Bolsa Família iniciado em 2004 no Brasil. Apesar de datar de 2004, o Programa Bolsa Família, que na continuidade deste trabalho será chamado somente de PBF, tem suas diretrizes baseadas no Programa Bolsa Escola de 2001.

Inicialmente adotado no âmbito municipal, o programa Bolsa Escola foi criado no município de Campinas (SP) e no Distrito Federal em 1995, e nos anos seguintes foi mimetizado em muitos outros municípios brasileiros. Devido ao sucesso obtido pelos municípios pioneiros no uso do programa, somado a popularidade dos programas de transferência condicionada de renda vigentes na América Latina, o

governo federal brasileiro decidiu expandir o programa Bolsa Escola para o âmbito federal em 2001. Este programa tinha como objetivo central incentivar que a totalidade das crianças completasse oito anos de estudo, ou seja, que as novas gerações conseguissem finalizar o ensino fundamental através de benefícios em dinheiro condicionados à frequência escolar. Em 2003, durante o primeiro ano do governo Luis Inácio Lula da Silva, o governo idealizou a expansão do programa Bolsa Escola. Além da expansão, o recém chegado governo também optou pela fusão de outros programas de assistência condicionada de renda que vigoravam no Brasil, como os programas Fome Zero, o Bolsa Alimentação e Auxílio Gás, em um programa maior e que atendesse um número maior de famílias, e ‘rebatizou’ o programa de Bolsa Família. Apesar de basear-se no programa Bolsa Escola, principalmente nas condicionalidades do recebimento da assistência, o PBF tem como foco principal o combate à pobreza.

Ativo desde 2004, o PBF concede benefícios para famílias em situação de pobreza ou pobreza extrema¹ compostas de crianças de 0 a 17 anos² ou gestantes. O benefício cedido varia de acordo com a composição familiar e a situação socioeconômica do domicílio. Até 2007, as famílias em extrema pobreza recebiam um benefício fixo de R\$58,00 mensais e um acréscimo de R\$18,00 para cada criança inscrita no programa, para no máximo três crianças por família, de forma a não incentivar um crescimento indesejado da natalidade. A partir de 2008, além da mudança na idade máxima, os benefícios aumentaram: as famílias em extrema pobreza passaram a receber um fixo de R\$62,00; o benefício por criança passou para R\$20,00; e o novo benefício por jovem de 16 ou 17 foi estabelecido em R\$30,00 mensais. Assim como seu antecessor, o PBF vincula o recebimento da renda do programa ao cumprimento de algumas condicionalidades por tipo de composição domiciliar:

- Para as famílias compostas por mulheres gestantes, o benefício é condicionado à realização de exames pré-natal;

¹ Até 2008, famílias com renda *per capita* familiar inferior a R\$120,00 e a R\$60,00 mensais, eram consideradas pobres e extremamente pobres respectivamente.

² Até 2008, a idade máxima das crianças era de 15 anos.

- para os domicílios com crianças de 0 a 6 anos, condicionado ao acompanhamento médico das vacinas, e a criança ter peso e altura adequados para sua idade; e
- para aqueles com crianças de 7 a 17, é necessária a matrícula regular no ensino fundamental assim como a frequência mínima de 85% do ano letivo.

Além de atender aos critérios de elegibilidade, para receberem o benefício do PBF, as famílias devem fazer a inscrição no programa na prefeitura de seu município, através do preenchimento de um questionário a respeito da composição do domicílio e da renda familiar e o recebimento do benefício é feito através de um cartão bancário entregue, geralmente, a mulher do chefe da família.

Após a ampliação do programa em 2004, o PBF se tornou uma importante ferramenta de combate à pobreza e a desigualdade de renda no Brasil. Em 2007, o programa atendeu mais de 11 milhões de famílias (aproximadamente um quarto da população brasileira) e em 2008 o orçamento gasto com o programa chegou a R\$12 bilhões. Assim, atualmente o PBF é o maior programa de transferência condicionada de renda do mundo, superando o programa mexicano *Oportunidades*. Mesmo com *status* de maior programa deste gênero, atualmente o governo federal destina cerca de 0,07% do PIB brasileiro ao programa, uma parcela pequena tendo em vista que o PBF atende aproximadamente 25% da população do país. É interessante citar também que, em geral, os benefícios sociais e de bem-estar gerados pelos programas de transferência condicionada de renda superam seus custos, como mostram Damon e Glewwe (2009) utilizando evidências de obtidas de diversos países latino-americanos e caribenhos.

Como colocado, o principal objetivo do PBF é atenuar a pobreza de curto-prazo, contribuindo, desta forma, na redução da desigualdade de renda. Vários estudos têm apresentado evidências empíricas de que o programa tem sido bem sucedido neste quesito. Soares *et. al.* (2006), por exemplo, estimam através dos dados da PNAD 2004 que 80% da renda alocada ao PBF foi destinada às famílias abaixo da linha da pobreza, e que o recebimento do benefício foi capaz de retirar parte destas famílias da miséria. Além disso, os autores encontram evidências que o

programa foi responsável por 21% da redução do Gini ocorrida entre 1995 e 2004³. Utilizando a mesma base de dados, Tavares *et. al.* (2009) também demonstram que o PBF é efetivo no combate a pobreza e a pobreza extrema, e que também se mostrou vital na recente redução da desigualdade de renda no país. Corroborando com estes resultados, Paes de Barros *et. al.* (2006) mostram que, mesmo em meio à estagnação econômica vivida pelo Brasil entre 2001 e 2004, houve uma importante redução da desigualdade, e apesar de não se mostrar o fator mais impactante, as transferências do governo exerceram papel importante nesta redução além de ter contribuído para a diminuição da pobreza. Outros trabalhos como Rocha (2005) e Ferreira *et. al.* (2006) também destacam a importância deste tipo de política na redução da pobreza e da desigualdade de renda no país.

Este tipo de política de transferência condicionada de renda não impacta, entretanto, somente de maneira “direta” a população. Resende e Oliveira (2008)⁴ mostram que famílias pobres beneficiadas por programas deste tipo têm seu consumo aumentado, e que a parcela do benefício alocada ao consumo é destinada à melhoria da dieta das famílias e à obtenção de itens relacionados à educação infantil, higiene e a saúde. Tais resultados sugerem que programas de transferência de renda aliviam a pobreza imediata dos domicílios e, além disso, podem gerar efeitos de longo prazo uma vez que boa parte dos recursos recebidos destina-se a melhoria do *status* nutricional, da educação e da saúde. Outro exemplo de impacto indireto do PBF é mostrado por Tavares (2008), que mostra que há a possibilidade de existir um incentivo adverso sobre a oferta de trabalho das mães de famílias beneficiárias. Seu trabalho concluiu que há uma redução de 10% da jornada de trabalho das mães de domicílios atendidos pelo programa. Ainda sobre a oferta de trabalho, Ferro, Kassouf e Levison (2009) investigam os efeitos da participação no programa antecessor ao PBF, o programa Bolsa-Escola de 2001, sobre as decisões de trabalho nos domicílios. Os autores concluem que a participação no programa reduz a oferta de trabalho das crianças e aumenta a oferta de trabalho dos pais das famílias beneficiadas.

Apesar de existir uma vasta bibliografia a respeito dos diversos impactos do PBF no Brasil, não são muitos os trabalhos que focam o impacto sobre indicadores

³ O índice Gini, que mede desigualdade, reduziu de 59.9 para 56.8 no período.

⁴ Utilizam a participação no Bolsa-Escola, já que utilizam a POF de 2002 à 2003.

educacionais. Utilizando dados municipais de 261 municípios situados em 5 estados nordestinos, coletados junto às secretarias destes municípios, Janvry, Finan e Sadoulet (2007) estimam que a participação no programa⁵ reduziu a taxa de abandono escolar em 7.8 pontos percentuais, entretanto elevou a taxa de reprovação em 0.8 pontos percentuais. Os autores argumentam que esse último resultado pode ser explicado pelo fato do benefício do programa ajudar a manter na escola um indivíduo com menor capacidade acadêmica ou com pouco interesse nos estudos e que, não fosse o benefício do programa, abandonaria a escola. Para obter estes resultados, o trabalho utiliza regressões em painel por município com dados entre 1999 e 2003, comparando alunos elegíveis com alunos que efetivamente recebem o benefício. Apesar dos resultados serem relevantes, este trabalho apresenta duas limitações na estimação do efeito do programa. Em primeiro lugar, mesmo utilizando somente alunos elegíveis ao programa assim como regressões em painel (neste caso as características não observáveis são computadas na estimação), este trabalho não utiliza outros controles para isolar o efeito do programa sobre o desempenho escolar como as características da escola, por exemplo. Isto significa que existe a possibilidade de que algumas características observáveis dos alunos, além da renda familiar *per capita*, que não foram incluídas no modelo e não são constantes no tempo, sejam correlacionadas com a participação no programa, o que pode gerar algum viés na estimação. Além disso, os autores utilizam dados de apenas 261 municípios brasileiros e todos pertencentes a 5 estados do Nordeste, ou seja, uma amostra não representativa do Brasil, e nem do Nordeste, o que de certa forma impossibilita a generalização do resultado encontrado.

Em um trabalho mais recente, Glewwe e Kassouf (2010) encontraram o impacto do PBF sobre o total de matrículas do ensino fundamental e sobre as taxas de abandono e aprovação através de dados do Censo Escolar de 1998 até 2005. Entre estes anos, o Censo Escolar continha um item, incluído no questionário respondido pela escola, com a seguinte pergunta: “algum aluno desta escola participa do PBF?” com as alternativas “sim” ou “não”. Desta forma, através do Censo Escolar destes anos era impossível identificar quais os alunos, assim como a

⁵ Os autores estimaram a participação no Bolsa-Escola, uma vez que seus dados são do período de 1999 a 2003.

quantidade de alunos beneficiados pelo programa em cada escola, de maneira que era possível saber somente se pelo um aluno da escola recebia o benefício. Utilizando regressões por escola, os autores mostram que escolas com pelo menos um aluno participante do programa tiveram um aumento médio de 6% nas matrículas, uma redução de aproximadamente 5 pontos percentuais da taxa de abandono e um aumento de 0.9 pontos percentuais na taxa de aprovação dos alunos da 1ª a 4ª série e de 0.3 pontos percentuais para os alunos da 5ª a 8ª série. Apesar de o trabalho utilizar um painel de 7 anos contendo todas as escolas públicas brasileiras, a base de dados utilizada na estimação destes resultados apresenta algumas limitações. Como mencionado, o Censo Escolar destes anos, assim como os Censos mais recentes⁶, carecia de informações de quantos alunos efetivamente recebiam o programa, de forma que os autores encontraram o efeito da disponibilidade do programa nas escolas uma vez que foram impossibilitados de encontrar o efeito do programa em si. Além disso, as regressões não utilizam características dos alunos como controle, de forma que é necessário supor que a composição dos alunos nestas escolas se mantenha constante ao longo do período utilizado. Está hipótese pode ser forte, uma vez que a criação do programa, que ocorreu dentro do período utilizado, por si só alterou a composição dos alunos devido ao choque positivo na renda familiar.

Com a finalidade de responder uma questão bastante semelhante a que será avaliada neste trabalho, Behrman, Segupta e Todd (2000) mostram que os alunos mexicanos que foram beneficiados pelo programa *Progresá* não apresentam diferenças estatisticamente significantes em relação aos não participantes do programa em testes de desempenho escolar. Outro exercício realizado pelos autores foi comparar o formato das distribuições das notas dos alunos beneficiados e não beneficiados e, novamente, não encontraram mudanças significativas nas distribuições do desempenho. Interessante notar que os autores destacam que estes resultados podem ser provenientes de limitações da base de dados utilizada⁷ e, além disto, é importante ressaltar que este resultado não pode ser generalizado para o caso brasileiro.

⁶ Inclusive, a partir de 2005 o Censo Escolar retirou de seu questionário a pergunta se algum aluno da escola recebeu benefícios do PBF.

⁷ A base utilizada pelos autores era de pouco tempo após a implementação do programa, e como os próprios argumentam, não houve tempo suficiente para o programa impactar sobre o desempenho.

Neste mesmo contexto, Pellegrina (2011) avalia o impacto do PBF sobre o desempenho escolar e a matrícula do alunado paulista. Utilizando as notas do SARESP de 2007 e 2009, além das notas de boletins escolares do período de 2007 a 2009 (coletadas junto a Secretaria da Educação do Estado de São Paulo, que também forneceu a informação de recebimento do auxílio do programa por parte dos alunos) o autor estimou os possíveis impactos do programa sobre diferentes variáveis associadas à educação, todos ao nível de indivíduo. Através de diferentes métodos de *matching* e de diferenças-em-diferenças, o autor encontrou que o programa tem efeito sobre as variáveis que estão diretamente atreladas as condicionalidades do PBF, entretanto não tem efeito sobre as variáveis de desempenho. O autor encontrou evidências de que a participação no PBF reduz o abandono e aumenta a freqüência escolar, entretanto parece não haver efeito sobre as notas em exames padronizados e sobre as notas do boletim escolar, mesmo quando feitos diversos testes de robustez. Apesar de utilizar um rico conjunto de controles e de considerar características não observáveis dos alunos em alguns de seus modelos, os resultados encontrados pelo autor também não podem ser generalizados para os demais estados brasileiros.

Alguns outros trabalhos discorrem sobre o impacto do PBF sobre o ensino brasileiro, como, por exemplo, o de Ferro, Kassouf e Levison (2009), que analisam também o impacto do programa sobre a freqüência escolar, uma vez que esta é uma das condições atreladas ao recebimento do benefício do programa. Apesar de suas preocupações focarem a oferta de trabalho infantil, os autores revelam que o recebimento do benefício do PBF aumenta a freqüência escolar.

Assim, fica evidente que existe certa carência de trabalhos que avaliem os possíveis impactos que este programa de transferência de renda pode trazer para a educação no Brasil. O presente trabalho busca contribuir nesta linha ao avaliar o impacto do PBF sobre o desempenho escolar médio das escolas, utilizando escolas do Brasil inteiro. Em um primeiro momento, deve-se esperar que o programa impacte de maneira positiva o desempenho escolar dos alunos por dois motivos: primeiro, o próprio programa impõem através de suas condicionalidades que os alunos beneficiados tenham presença mínima de 85% nas aulas, superior ao normalmente exigido (em geral, as escolas exigem presença mínima de 75% nas aulas) – se freqüência às aulas facilita o aprendizado, espera-se que esta exigência contribua para o aumento do desempenho escolar. Em segundo lugar, pelo efeito do

aumento de renda, que indiretamente afeta de forma positiva o desempenho escolar. Os canais pelos quais a renda familiar afeta o desempenho escolar são muitos, por exemplo, através da melhora do *status* nutricional das crianças (Resende e Oliveira 2008). Assim, percebe-se que o PBF, individualmente, deve impactar de forma positiva sobre o desempenho dos alunos, entretanto este resultado pode ser diferente do ponto de vista agregado. Isto é, como argumentam Janvrym Finan e Saudolet (2007), há a possibilidade de que indivíduos desinteressados ou com menor capacidade acadêmica se mantenham matriculados na escola, estimulados somente pelo benefício do programa, o que deve reduzir o desempenho médio das escolas com um número maior de alunos participantes do programa. Desta forma, este trabalho tem como objetivo estudar se mudanças na proporção de alunos que recebem o benefício do PBF alteram o desempenho médio da escola.

O presente trabalho combina os dados obtidos através dos “Microdados do Acompanhamento da Frequência do Programa Bolsa Família”⁸ 2008, Censo Escolar de 2008 e da Prova Brasil 2009, procura analisar os possíveis impactos que o recebimento do benefício do Programa Bolsa Família pode gerar sobre o desempenho escolar médio dos alunos da quarta série do ensino fundamental. A continuidade do trabalho é dividida da seguinte forma: na seção seguinte há uma descrição da origem dos dados, assim como uma análise descritiva dos mesmos; a 3ª seção trata da metodologia empregada para obter o efeito do programa; a 4ª apresenta e analisa os resultados obtidos e finalmente a última seção conclui o trabalho.

⁸ Esta base consiste no sistema de condicionalidades do PBF, que é um banco de dados que reúne informações das famílias beneficiadas pelo programa assim como o cumprimento das condições impostas pelo programa.

2. Descrição das Bases de Dados utilizadas e Análise descritiva dos dados obtidos

Como mencionado anteriormente este trabalho utilizará os dados do Censo Escolar de 2008. Trata-se de um levantamento de dados sobre matrícula e movimentação dos alunos, bem das condições de oferta de ensino, em todos os estabelecimentos educacionais do país. É realizado anualmente pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), autarquia do MEC (Ministério da Educação). As informações são coletadas em dois momentos distintos: no início de cada ano letivo, há uma coleta que visa apurar a situação de rendimento dos alunos relativa ao ano anterior (grosso modo: aprovação, reprovação, abandono ou transferência); num segundo momento, especificamente no final de abril, são coletadas as informações de matrícula do ano letivo corrente. Todas as informações são obtidas através do preenchimento de um questionário via internet, feita por um representante de cada escola. Os dados se encontram no portal do INEP.

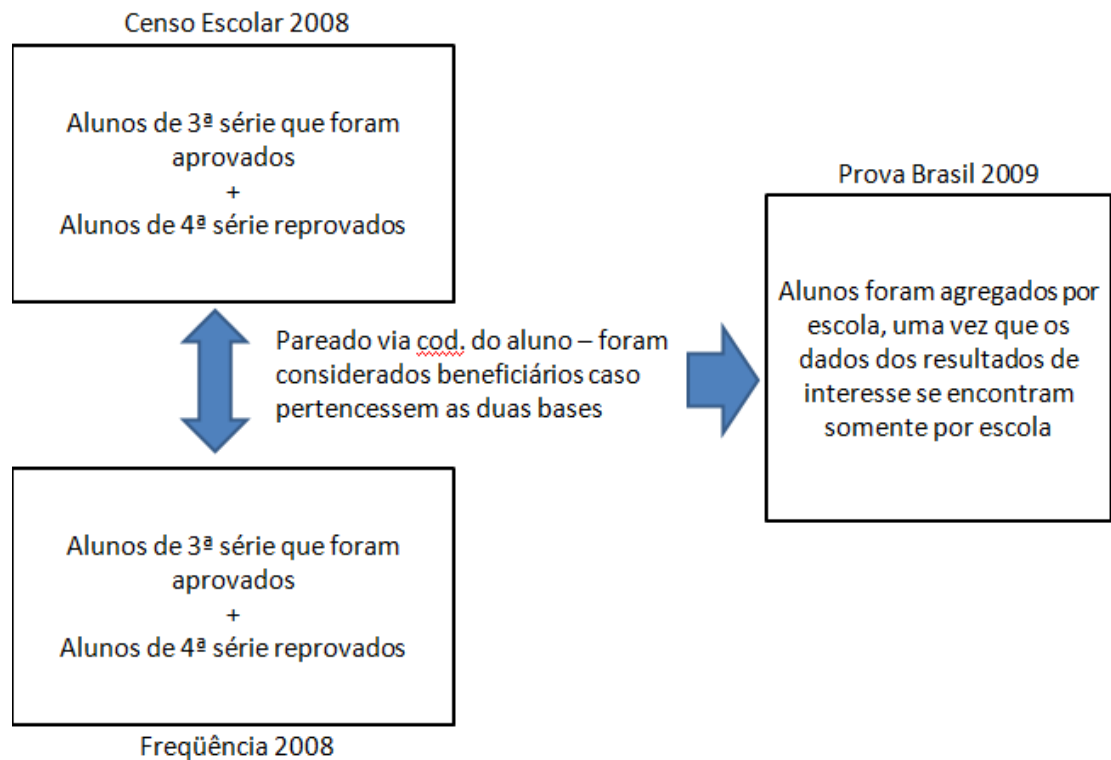
Como já colocado, no Censo Escolar não consta informação sobre o recebimento de benefícios de programas de transferência de renda por parte dos alunos. Assim, utiliza-se outro conjunto de dados, especificamente, os dados de acompanhamento da frequência escolar dos beneficiários do Programa Bolsa Família, coletado pela SECAD/MEC. Estes dados foram ‘casados’ com os dados do Censo Escolar, com base em diversas informações dos alunos como o próprio nome do beneficiário, o número de Identificação Social (NIS), data de nascimento, sexo e o código da Família do beneficiário no CADÚNICO. Nem todos os alunos presentes no cadastro de acompanhamento do Bolsa Família foram ‘encontrados’ no Censo Escolar – no entanto, para este trabalho todos os alunos que aparecem no Censo Escolar mas não aparecem no Cadastro de Acompanhamento foram definidos como ‘não- beneficiários’ do PBF. Isso significa que alunos beneficiários podem ter sido erroneamente classificados como não-beneficiários, o que levaria a uma subestimação do impacto do programa.

Além destes bancos, este trabalho utiliza também a base de dados da Prova Brasil de 2009 para avaliar o desempenho escolar dos indivíduos. Criada em 2005, a

Prova Brasil é um teste realizado em todas as escolas brasileiras para a 4ª e a 8ª série do ensino fundamental a cada dois anos. Realizada no mês de Novembro dos anos ímpares, esta avaliação tem por objetivo medir a proficiência (nível de aprendizagem) dos alunos brasileiros. Consiste em uma prova de português para avaliar o foco em leitura, e uma de matemática para avaliar a capacidade de resolução de problemas. A junção dos dados da Prova Brasil com os dados do Censo Escolar é feita pelo código da escola – ou seja, não foi possível parear ao nível do aluno os dados ‘censo escolar + acompanhamento do Bolsa Família’ com os dados de rendimento da Prova Brasil, uma vez que ainda não estão disponíveis os microdados da Prova Brasil 2009, contendo o desempenho por aluno.

Um ponto importante sobre a disponibilidade dos dados. O cruzamento dos dados do Censo Escolar e Cadastro de Acompanhamento do Bolsa Família foi feito em parceria pelas equipes do MEC e MDS especificamente para o ano de 2008, ano em que a Prova Brasil não foi realizada. Desta forma, este trabalho utiliza as informações da “provável” 4ª série de 2009, ou seja, os alunos que estavam na 3ª série em 2008 e que foram aprovados e os alunos da 4ª série de 2008 e que foram reprovados neste ano para explicar o desempenho dos alunos da 4ª série contido na Prova Brasil 2009. Apesar de não utilizar rigorosamente os mesmos alunos, visto que alguns alunos podem ter evadido a escola, é pouco provável que entre estes dois anos tenha ocorrido alguma mudança tão abrupta na transição destes alunos entre os anos. Assim, será feita uma hipótese de que as escolas não se alteram de um ano para o outro, e que os alunos que estavam na 3ª série em 2008 serão os mesmos analisados na 4ª série de 2009, ou de uma forma menos restritiva, os alunos não precisam ser os mesmos, mas suas características sim.

Resumidamente, este trabalho utilizará informações dos alunos da 3ª série de 2008 para avaliar o impacto da participação no PBF neste ano sobre a taxa de aprovação e abandono média das escolas em 2009 assim como sobre a proficiência média das escolas no exame da Prova Brasil para os alunos da 4ª série deste mesmo ano, assumindo que a composição dos alunos nas escolas se manterá semelhante entre 2008 e 2009. O diagrama a seguir explicita melhor como foram pareadas as bases de dados



Como mencionado no início desta seção, devido a um possível pareamento imperfeito das bases utilizadas, há possibilidade de que se subestime o número de alunos participantes do programa. Entretanto, como será apresentada a seguir, a proporção de indivíduos beneficiados pelo programa na amostra parece ser bastante condizente com a realidade, indicando que esta limitação da base de dados possivelmente não viesará os resultados. A tabela 1 mostra a proporção de alunos que recebem o benefício do PBF.

Tabela 1: Proporção de alunos que recebem benefícios do PBF

Proporção de alunos beneficiados na amostra				
	Freqüência	Porcentagem		
Participam do PBF	527269	25,52%		
Não participantes	1538525	74,48%		
Total	2065794	100,00%		
Média de beneficiados nas escolas				
No de escolas	Proporção média	DP	Min	Max
29503	29,33%	0,4552	0	100,00%

Fonte: Censo Escolar 2008 e Freqüência 2008

Os resultados apresentados na tabela 1 mostram que a porcentagem de alunos tratados na amostra se aproxima de 25%, número bastante realista uma vez que um quarto da população brasileira efetivamente participa do programa. A tabela 2 reporta o desempenho médio e a taxa de aprovação média nas escolas:

Tabela 2: Desempenho dos alunos por escola em 2009

Por escola:	Observações	Média	DP	Min	Max
Taxa de aprovação	28845	88,43%	11,23	0	100%
Taxa de abandono	28845	1,90%	4,27	0	100%
Proficiência em Mat.	27088	201,98	24,24	121,21	319,26
Proficiência em Pt.	27088	181,93	20,3	116,24	292,49

Fonte: Censo Escolar 2008, Freqüência 2008 e Prova Brasil 2009

A tabela 2 revela que a média da taxa de aprovação, da taxa de abandono e exibe as notas médias das escolas nos exames de proficiência da Prova Brasil. A tabela 3 apresenta a correlação entre a proporção de alunos participantes do programa nas escolas e os resultados de interesse.

Tabela 3: Correlação entre o desempenho escolar e a proporção de beneficiados

	Proporção de alunos beneficiados pelo PBF
Taxa de aprovação	-0,2436
Taxa de abandono	0,1810
Proficiência em Mat.	-0,3855
Proficiência em Pt.	-0,4077

Fonte: Censo Escolar e Freqüência 2008 e Prova Brasil 2009

Nota-se que a participação no tratamento parece se relacionar negativamente com o desempenho escolar, uma vez que as correlações têm sinal negativo e são estatisticamente relevantes a 1%. Neste ponto é importante salientar que os dados apresentados na tabela 3 não demonstram que o recebimento da assistência do PBF impacta negativamente nos resultados de interesse, mostram apenas que em média, escolas com uma proporção maior de alunos tratados têm taxas de aprovação menor, taxa de abandono maior, e nota médias nas provas de proficiência também menores, não necessariamente havendo causalidade entre a participação no programa e o desempenho dos alunos. A tabela 4 apresenta as características médias dos alunos nas escolas.

Tabela 4: Composição dos alunos nas escolas

	Média	DP	Min	Max
No de alunos na escola	128,35	95,04	11	1464
Média de Idade	9,98	0,57	9	19,2
Tamanho da turma	26,19	6,77	1	96
Proporção de meninas	0,46	0,09	0	1
Proporção de negros e pardos	0,23	0,23	0	1
Proporção de brancos	0,18	0,23	0	1
Proporção de índios	0,002	0,02	0	1
Proporção de asiáticos	0,005	0,02	0	0,94
Prop. Não responderam etnia	0,58	0,49	0	1,00
Prop. de mães c/EF	0,13	0,06	0	0,64
Prop. de mães c/EM	0,100	0,070	0	0,52
Prop. de mães c/ES	0,070	0,06	0	1
Prop. de alunos não responderam sobre a escol. da mãe	0,266	0,11	0	1
Proporção de alunos que moram em região rural	0,22	0,37	0	1
Observações	29503			

Fonte: Censo Escolar 2008 e Prova Brasil 2007

A respeito da tabela 4, é interessante mencionar três fatos sobre a amostra utilizada. Primeiramente, no Censo Escolar a etnia é reportada pelo próprio respondente, como a maioria das pesquisas demográficas. Isto é, a proporção de alunos brancos na escola retrata na verdade a proporção de alunos que se declararam brancos, de modo que diversos alunos brancos que nada declararam a respeito de sua etnia não são contabilizados nesta variável. Desta forma, é possível que as variáveis de etnia não tenham seus impactos devidamente estimados. Outro ponto a se dizer a respeito da amostra é que os dados utilizados para encontrar a média da escolaridade das mães nas escolas (assim como os demais indicadores socioeconômicos médios das escolas) foram obtidos através do questionário da Prova Brasil 2007, uma vez que os microdados da Prova Brasil 2009 ainda não estão disponíveis. Interessante salientar também que o questionário da Prova Brasil a respeito da escolaridade da mãe dos alunos foi respondido pelos próprios alunos, havendo uma opção “não sei a escolaridade da minha mãe”. Novamente, devido à possibilidade de diversos alunos não terem respondido sobre a escolaridade de sua

mãe, é possível que as variáveis referentes a escolaridade das mães dos alunos nas escolas tenham seus impactos subestimados também. Finalmente, é importante ressaltar que na Prova Brasil de 2009 as escolas rurais foram incluídas no teste, de forma que é possível incluir escolas rurais na estimação.

Similarmente a tabela 4, a tabela 5 apresenta as características médias dos alunos, separando por janelas de escolas, ou seja, as escolas são separadas em quatro grupos: as que têm menos de 25% de alunos assistidos pelo PBF, as que têm entre 25% e 50% dos seus alunos participantes do programa, as que têm entre 50% e 75% e, finalmente, escolas com a proporção de alunos beneficiados maior que 75%.

Tabela 5: Características das Escolas por quartil de Proporção de beneficiados

	PBF < 25%	25% <PBF< 50%	50% <PBF< 75%	PBF > 75%
	Média /Desvio Padrão			
Taxa média de aprovação	91,56 (9,24)	87,11 (11,55)	85,70 (12,59)	82,41 (16,43)
Taxa média de abandono	1,05 (2,61)	2,43 (4,15)	2,36 (4,35)	5,41 (10,53)
Nota em Mat	211,64 (22,18)	195,25 (22,76)	189,88 (23,52)	185,64 (20,74)
Nota em Pt	190,23 (18,70)	175,95 (18,73)	170,47 (18,93)	167,11 (17,59)
Tamanho da escola	165,34 (112,99)	130,22 (82,52)	106,36 (72,37)	62,29 (51,47)
Média de idade	9,75 (0,54)	10,18 (0,522)	10,26 (0,516)	11,68 (1,47)
% Menina	0,474 (0,076)	0,4611 (0,082)	0,451 (0,093)	0,339 (0,315)
% Negros e Pardos	0,211 (0,202)	0,232 (0,242)	0,227 (0,250)	0,261 (0,347)
% Brancos	0,29 (0,272)	0,122 (0,188)	0,091 (0,148)	0,070 (0,192)
% Índios	0,001 (0,015)	0,001 (0,017)	0,000 (0,006)	0,004 (0,048)
% Asiáticos	0,004 (0,03)	0,004 (0,033)	0,002 (0,019)	0,009 (0,082)
% não responderam etnia	0,49 (0,482)	0,63 (0,482)	0,67 (0,468)	0,65 (0,478)
% Mor. Rurais	0,062 (0,155)	0,094 (0,195)	0,138 (0,246)	0,153 (0,304)

Escola Rural	0,006 (0,083)	0,013 (0,117)	0,0279 (0,164)	0,036 (0,188)
Escola Federal	0,001 (0,032)	0,000 (0,008)	0,000 (0,000)	0,001 (0,044)
Escola Estadual	0,384 (0,486)	0,275 (0,446)	0,291 (0,454)	0,494 (0,500)
Tamanho médio da turma	27,54 (6,16)	26,60 (6,53)	24,59 (6,59)	26,90 (8,64)
Água filtrada	0,842 (0,363)	0,861 (0,345)	0,897 (0,303)	0,903 (0,296)
Não possui energia	0,000 (0,029)	0,000 (0,016)	0,000 (0,000)	0,000 (0,000)
Sala de diretor	0,934 (0,248)	0,901 (0,297)	0,878 (0,326)	0,881 (0,323)
Sala de professores	0,694 (0,460)	0,494 (0,499)	0,473 (0,499)	0,470 (0,499)
Lab. Informática	0,608 (0,488)	0,402 (0,490)	0,330 (0,470)	0,498 (0,500)
Lab. Ciências	0,145 (0,352)	0,087 (0,281)	0,066 (0,249)	0,096 (0,296)
Atend. Especial	0,218 (0,413)	0,146 (0,353)	0,082 (0,275)	0,096 (0,296)
Quadra	0,646 (0,478)	0,457 (0,498)	0,380 (0,485)	0,480 (0,500)
Cozinha	0,975 (0,155)	0,969 (0,171)	0,973 (0,159)	0,957 (0,202)
Biblioteca	0,740 (0,438)	0,633 (0,481)	0,612 (0,487)	0,664 (0,472)
Parquinho	0,253 (0,434)	0,159 (0,366)	0,105 (0,307)	0,029 (0,168)
Número de Salas	11,23 (5,87)	10,23 (5,78)	9,31 (5,39)	9,86 (5,01)
TV	0,986 (0,116)	0,973 (0,161)	0,970 (0,170)	0,951 (0,214)
VHS	0,857 (0,349)	0,747 (0,434)	0,734 (0,441)	0,784 (0,411)
DVD	0,960 (0,194)	0,927 (0,259)	0,912 (0,282)	0,908 (0,288)
Número de Comp.	11,56 (11,02)	6,87 (8,02)	5,80 (23,04)	11,12 (49,36)
Internet	0,762 (0,425)	0,496 (0,500)	0,212 (0,408)	0,515 (0,500)
Banda Larga	0,567 (0,495)	0,264 (0,441)	0,212 (0,408)	0,277 (0,448)
% mães de alunos com EF completo	0,140	0,125	0,109	0,126

	(0,056)	(0,063)	(0,065)	(0,060)
% mães de alunos com EM completo	0,126	0,084	0,062	0,101
	(0,073)	(0,060)	(0,053)	(0,073)
% mães de alunos com ES completo	0,094	0,067	0,053	0,055
	(0,064)	(0,051)	(0,046)	(0,041)
% mães de alunos sem EF completo	0,37	0,45	0,50	0,47
	(0,014)	(0,165)	(0,173)	(0,147)
% de alunos que não responderam sobre a esc. da mãe	0,26	0,26	0,26	0,25
	(0,102)	(0,118)	(0,128)	(0,107)
Observações	12564	13050	3410	479

Fonte: Censo Escolar 2008, Freqüência 2008 e Prova Brasil 2007/2009

Assim como os resultados apresentados nas tabelas 2 e 3, os resultados aqui reportados mostram que a proporção de alunos beneficiários na escola é negativamente correlacionada com a proficiência dos alunos, tanto em matemática quanto em português, de forma que à medida em que se aumenta a proporção de alunos beneficiados nas escolas, a média da nota nos exames de proficiência se reduz. De forma análoga, escolas com porcentagem de beneficiários maior, em média tem taxa de aprovação menor e taxa de abandono maior. Outro ponto interessante a se salientar é que escolas com proporção maior de beneficiados têm, na média, infra-estrutura pior, são escolas menores, tem uma proporção menor de mãe com maior grau de escolaridade, e há uma chance bem maior da escola se localizar em zona rural.

Outra característica interessante de se investigar na amostra, é a localização das escolas por região, isto é, onde estão localizadas as escolas com maior número de alunos beneficiários. A tabela 6 mostra a distribuição das escolas, dividida pela proporção de alunos tratados pelo programa, por região.

Tabela 6: Proporção de escolas em cada quartil por região

Região	PBF<25%	Entre 25 e 50%	Entre 50 e 75%	PBF>75%	Total
Sul	3.799	2.207	295	18	6.319
	60,12%	34,93%	4,67%	0,28%	
Sudeste	6.942	4.054	968	133	12.097
	57,39%	33,51%	8,00%	1,10%	
Centro-Oeste	1.658	1.479	112	33	3.228

	51,36%	45,82%	3,47%	1,02%	
Norte	907	1.424	178	34	2.543
	35,67%	56,00%	7,00%	1,34%	
Nordeste	1.458	5.955	1.983	298	9.694
	15,04%	61,43%	20,46%	3,07%	
Total	14.764	15.119	3.536	516	33.935

Fonte: Censo Escolar 2008 e Freqüência 2008

Como esperado, os valores apresentados na tabela 6 retratam que as regiões mais pobres do Brasil comportam, proporcionalmente, uma maior quantidade de escolas com número elevado de alunos beneficiados pelo PBF. Este fato fica evidente quando se observa que mais de 75% das escolas na região Nordeste possui mais de um quarto de seus alunos pertencentes ao programa de transferência de renda, e na região Norte mais de 65% das escolas detém esta característica. Este número é bastante diferente nas demais regiões do país, onde menos da metade das escolas comporta mais de 25% de alunos beneficiados.

3. Metodologia

Sejam a aprovação, abandono e as notas nos exames de matemática e português os resultados de interesse do aluno i da 4ª série do Ensino Fundamental que estuda na escola s , todos denotados por Y_{is} . Como pode ser observado em Menezes Filho (2007), na literatura há certo consenso de que o desempenho escolar dos alunos depende das características da escola a qual o aluno atende, dos docentes desta escola e, principalmente das características socioeconômicas e de *background* familiar do aluno. Em um primeiro momento, admita que Y_{is} seja uma função linear da probabilidade do aluno participar do PBF, que condensa as características socioeconômicas e familiares do aluno, e de uma variável binária b_{is} que será igual a ‘um’ quando o indivíduo efetivamente receber o benefício do programa e ‘zero’ caso contrário. Formalmente, Y_{is} será dado pelo seguinte processo:

$$Y_{is} = \alpha + \beta \cdot P_{is}(c_{is}) + \phi \cdot b_{is} + \varepsilon_{is} \quad (1)$$

, onde ε_{is} é o termo de erro aleatório.

Desta forma, considere Y_{is}^0 o resultado para o aluno que não recebeu o benefício do programa, já que $b_{is} = 0$, e Y_{is}^1 o resultado para o aluno que recebeu. Idealmente, se fosse possível observar o mesmo indivíduo na situação em que ele fosse tratado simultaneamente à situação em que ele não fosse tratado, a subtração de $Y_{is}^1 - Y_{is}^0$, geraria o efeito do programa para este indivíduo, e a média destes efeitos para todos os indivíduos geraria o ATE (efeito médio do tratamento). Contudo é impossível observar o mesmo indivíduo nas duas situações: ou ele recebe ou ele não recebe. Assim, faz-se necessário comparar o resultado de alunos distintos. Caso a probabilidade do aluno participar do programa fosse conhecida, seria possível observar alunos com a mesma probabilidade de ser beneficiário do PBF recebendo e não recebendo (ou seja, admitindo que houvesse algum erro de focalização), e a comparação do resultado entre esses alunos geraria uma boa estimativa do efeito do tratamento.

Entretanto, a probabilidade de participar do PBF não é conhecida. Assim, é necessário obter uma estimativa para esta probabilidade. Admita que a probabilidade de participar do PBF possa ser descrita por uma função não linear das características dos alunos, c_{is} , que engloba suas características socioeconômicas, uma vez que a participação no PBF depende primordialmente da renda familiar *per capita*, e geográficas. Isto é, como o desenho do programa prioriza municípios mais pobres, portanto é simples imaginar que uma família pobre que reside em um município com renda per capita baixa tenha maiores chances de receber o benefício de que uma família que habita um município mais rico. Assim, uma estimativa desta probabilidade seria dada por:

$$P_{is}^{\wedge}(c_{is}) = f(\delta + \gamma \cdot c_{is}) \quad (2)$$

, assumindo que $f(.)$ tem distribuição logística, de forma que a equação (2) possa ser estimada por um modelo Logit.

Neste ponto, se faz necessário assumir que as estimativas de $P_{is}^{\wedge}(c_{is})$ obtidas pela equação (2) sejam suficientemente precisas, ou seja, esta probabilidade está considerando características observáveis suficientes para que a seguinte condição seja atendida:

$$Y_{is}^1, Y_{is}^0 \perp b_{is} / P_{is} \quad (3)$$

Em outras palavras, a condição dada por (3) significa que dadas às características do indivíduo, condensadas em uma probabilidade, a designação ao tratamento é randômico, ou mais claramente, se for garantido que os alunos sejam efetivamente semelhantes, o fato de um receber o auxílio do PBF e seu par não receber é puramente aleatório. Além disso, é necessário também assumir que:

$$0 < P_{is}(\cdot) < 1 \quad (4)$$

, que significa que na amostra utilizada, existem alunos semelhantes tratados e não tratados suficientes para que o efeito do programa seja obtido, condição esta que é

comumente chamada na literatura de avaliação de políticas públicas de suporte comum. Dadas as condições (3) e (4), podemos estimar a equação (1) através de:

$$Y_{is} = \alpha + \beta \cdot P_{is}^{\wedge}(c_{is}) + \phi \cdot b_{is} + \varepsilon_{is} \quad (1')$$

, onde $P_{is}^{\wedge}(c_{is})$ denota as estimativas obtidas na estimação de (2), de maneira o efeito do PBF seja dado pelo parâmetro ϕ . Devido à indisponibilidade dos dados a respeito do resultado por aluno, é possível, sem perda de generalidade do modelo, utilizar uma regressão por escola. Assim, utilizando as mesmas hipóteses, pode-se alternativamente estimar a seguinte equação:

$$Y_s^* = \mu + \eta \cdot P_s^{\wedge*}(\cdot) + \varphi \cdot b_s^* + \xi_s \quad (1'')$$

, onde $Y_s^* = \frac{\sum_1^{N_s} Y_{is}}{N_s}$, $P_s^{\wedge*}(\cdot) = \frac{\sum_1^{N_s} P_{is}^{\wedge}(\cdot)}{N_s}$, $b_s^* = \frac{\sum_1^{N_s} b_{is}}{N_s}$, N_s é o número de alunos na escola s e ξ_s denota o termo de erro. Devido as alterações nas variáveis da equação oriundas da transformação de (1') para (1''), tanto os coeficientes, quanto as próprias variáveis, tem seu significado alterado. Isto é, Y_s^* passa a denotar o resultado médio para a escola, $P_s^{\wedge*}(\cdot)$ a probabilidade média de um aluno desta escola participar do PBF e b_s^* a proporção efetiva de alunos que recebem o tratamento. Quanto aos coeficientes, é importante ressaltar que φ agora denota o efeito marginal de um aumento na proporção de alunos participantes do programa sobre o desempenho médio da escola. Se atendidas as hipóteses de identificação (3) e (4), as estimativas obtidas por OLS de (1'') serão consistentes.

Resumidamente, este trabalho fará uma estimação em 2 estágios, sendo o primeiro a estimação da probabilidade dos alunos receberem o benefício do programa dada pela equação (2) através de um Logit, e o segundo, dadas estas probabilidades, a estimação de uma equação por OLS da equação por escola dada por (1''), assumindo que as hipóteses de que, dadas as características dos alunos, a atribuição ao tratamento é aleatória, e de que, para cada aluno tratado, tenha um aluno não tratado com probabilidade de receber o tratamento semelhante.

Note que a estimação do resultado de interesse, tanto por aluno, quanto por escola não utiliza características da escola na obtenção do parâmetro de interesse. Como mencionado no início desta seção, apesar de apresentarem causalidade com

o desempenho escolar, as características físicas das escolas foram excluídas do modelo propositalmente em um primeiro momento com o intuito de testar a robustez do modelo. Observe que na equação (1), assim como em suas transformações subseqüentes (1') e (1''), as características da escola são englobadas no erro da equação, uma vez que estas determinam parte do desempenho escolar dos alunos. Apesar de correlacionadas com os resultados de interesse, espera-se que as características da escola, condicional na probabilidade de ser beneficiário do PBF, não sejam correlacionadas com o fato de os alunos receberem ou não o auxílio do programa. Isto é, espera-se que para escolas com alunos de mesmas características socioeconômicas (denotadas por $P_s^{\wedge*}(\cdot)$), o fato de uma escola ter mais alunos tratados do que a outra não está correlacionado com as características destas escolas. Assim, tem-se que a condição $E(b_s^*, \xi_s) = 0$ possa ser garantida em (1''), tornando as estimativas de OLS para φ consistentes.

No mesmo contexto, considere agora que as características da escola sejam denotadas por S_s para a escola s . Ao incluir estas características na equação (1''), têm-se:

$$Y_s^* = \zeta + \theta \cdot P_s^{\wedge*}(\cdot) + \varphi \cdot b_s^* + \lambda \cdot S_s + \psi_s \quad (5)$$

Desta maneira, o objetivo de estimar (1'') e (5) em dois exercícios separados é verificar se há alteração na magnitude do coeficiente φ estimado pelas duas equações e, por conseqüência, observar se a condição $E(b_s^*, \xi_s) = 0$ está sendo suprida em (1''). Caso se observe alguma discrepância significativa em φ nos dois modelos, pode-se constatar que as estimativas de OLS para φ obtidas através de (1'') são inconsistentes devido a um problema de viés de variável omitida. Isto significa que as estimativas obtidas pela equação (2) não foram suficientemente precisas, assim alguma característica que não foi computada na estimação de $P_s^{\wedge*}(\cdot)$, que é correlacionada com o recebimento do benefício do programa, foi incorporada pelo termo de erro ξ_s , resultando em $E(b_s^*, \xi_s) \neq 0$. Assim, a inclusão de variáveis referentes à escola deve, por alguma forma, capturar parte deste efeito ignorado pelas estimativas obtidas em (2), o que pode alterar significativamente φ , tornando sua estimativa viciada.

Para exemplificar a situação descrita acima, imagine um caso hipotético em que a escolaridade das mães dos alunos, uma variável que captura tanto a renda como o *background* da família, não fosse incluída na estimação da probabilidade do aluno receber o benefício do PBF. Neste caso, esta variável seria incorporada no termo de erro ξ_s , já não fora incorporada em $P_s^{\wedge*}(\cdot)$ e ela é claramente correlacionada com o desempenho escolar. Note que por ser bastante correlacionada com a renda familiar, a escolaridade das mães provavelmente determina a participação ou não no programa. Assim, por ser relacionada com b_s^* , a ausência da variável escolaridade das mães na estimação de $P_s^{\wedge*}(\cdot)$, levaria a $E(b_s^*, \xi_s) \neq 0$, já que esta fora incorporada no termo de erro em (1''). Em outras palavras, a imprecisão das estimativas de (2) impedem que a designação ao tratamento seja aleatória na estimação de (1'') que, como visto na condição (3), é de suma importância para a estimação do efeito do programa.

4. Resultados

4.1. Estimação da probabilidade do aluno receber o tratamento

A seguir, serão apresentados os resultados obtidos para as estimações das probabilidades individuais de receber assistência do PBF, dadas as características do aluno. A estimação, feita através de um modelo Logit, utiliza como variáveis dependentes o sexo, a etnia, se o aluno mora em zona rural, a UF que ele reside, se ele utiliza transporte público, o tamanho do município onde reside⁹, a proporção de famílias com renda *per capita* inferior à R\$120,00¹⁰ no município e a escolaridade média das mães das escolas onde os alunos estudam (utilizando cluster de escola). Como antecipado na seção anterior, a participação no programa depende basicamente da renda e da localização geográfica da família e, devido à indisponibilidade da informação sobre a renda familiar dos alunos, a estimação da probabilidade dependerá das variáveis mencionadas acima já que estas devem capturar grande parte do efeito da renda assim como o efeito geográfico. A tabela 7 apresenta o resultado destas estimações:

Tabela 7: Estimação da probabilidade do aluno participar do PBF

	Coeficiente/Std. Err.	Efeito Marginal
Menina	0,0442* (0,0034)	0,0076
Negro/Pardo	0,2366* (0,0075)	0,0424
Não respondeu sobre sua etnia	0,3421* (0,0087)	0,0589
Indígena	0,1068** (0,0527)	0,0190
Asiático	0,1618* (0,0375)	0,0293
Mora em zona rural	0,0491* (0,0133)	0,0086

⁹ Dado pelo número total de habitantes, obtido através do Censo Demográfico 2010 do IBGE.

¹⁰ Dado proveniente do Censo Demográfico 2010 do IBGE.

Tamanho do município	0,0005*** (0,0003)	0,0001
Proporção de famílias com renda <i>per capita</i> <R\$120,00	1,2128* (0,0478)	0,2103
Transp. Público	0,0937* (0,0120)	0,0165
% Mães com EF nas escolas	-0,1678** (0,0725)	-0,0291
% Mães com EM nas escolas	-1,8882* (0,0731)	-0,3274
% Mães com ES nas escolas	-2,5245* (0,0828)	-0,4378
% Mães com esc. Desconhecida	-0,072* (0,0179)	-0,0092
Dummys de UF	SIM	
Observações	2065794	
Pseudo - R2	0,0882	

*p<0,01; **p<0,05 e ***p<0,10

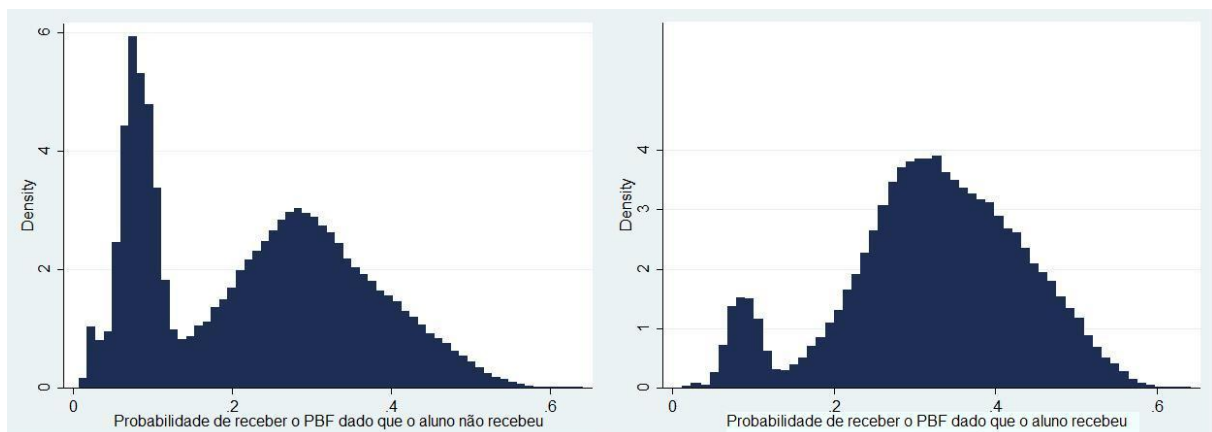
Os resultados reportados na tabela 7 são bastante intuitivos, isto é, os coeficientes apresentam seus sinais esperados. Através destas estimativas foi atribuída uma probabilidade de cada indivíduo da amostra receber o benefício do programa, dadas suas características pessoais, como antecipado na 4ª seção deste trabalho. Com a finalidade de investigar a capacidade preditiva do modelo, a seguir na tabela 8 serão reportadas as estimativas de recebimento do benefício do programa *versus* o recebimento efetivo do programa, considerando que se estimou que um aluno fosse tratado caso sua probabilidade de receber o benefício superasse 0.25 (fosse maior que a média da população brasileira).

Tabela 8: Estimativa de recebimento x Recebimento efetivo

Estima-se que recebeu	Recebeu efetivamente		
	NÃO	SIM	Total
NÃO	803759 38,91%	116358 5,63%	920117 44,54%
SIM	734766 35,57%	410911 19,89%	1145677 55,46%
Total	1538525 74,48%	527269 25,52%	2065794 100%

A tabela 8 apresenta algumas características interessantes a respeito das probabilidades estimadas. Primeiramente, o modelo prevê que cerca de 55% da população deveria receber o benefício, número aparentemente realista, uma vez que o PBF mesmo sendo direcionado a maior parte a população do país (a parcela pobre e muito pobre no Brasil é superior a metade da população), o programa ainda não logrou êxito em atingir a totalidade desta população. Assim, o fato de se observar somente uma parte da população elegível receber de fato o benefício (no caso, 55% são estimados como elegíveis, ou seja, deveriam receber o benefício, enquanto 25% recebem efetivamente) parece adequado. Outro ponto interessante de se observar, é que cerca de 80% dos alunos efetivamente beneficiados foram identificados como tal pelo modelo. Ademais, o erro de focalização previsto pelo modelo, ou seja, a parcela da população que não deveria receber o tratamento e recebe, corresponde a 5% da população. Prosseguindo com a análise destas variáveis, o gráfico 1 apresenta histogramas das probabilidades estimadas para os alunos que receberam o tratamento e para os que não receberam.

Gráfico 1: Probabilidade de receber o benefício para tratados e não tratados



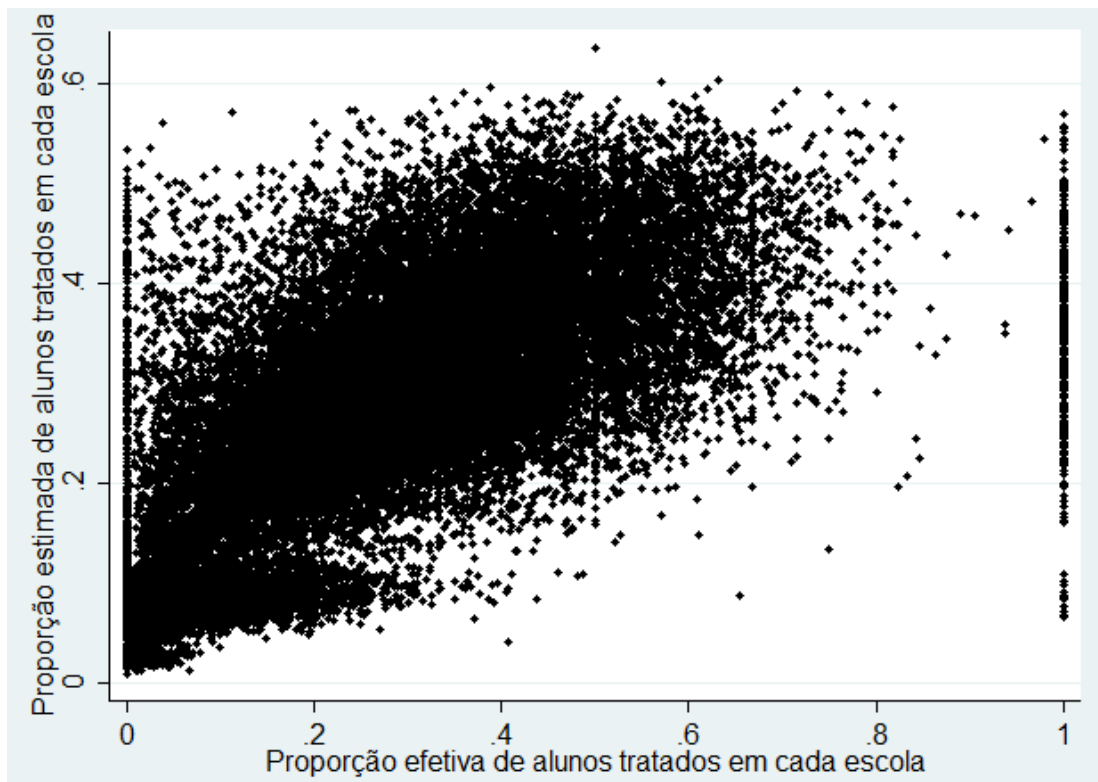
Através do gráfico 1 pode-se observar algumas características interessantes a respeito dos grupos. Além de este gráfico mostrar que, como esperado, o grupo tratado tem uma probabilidade média de receber o benefício maior do que o grupo de controle, ele apresenta algo importante para o *framework* deste trabalho, isto é, o gráfico mostra que para cada probabilidade de receber o tratamento em geral existe um indivíduo efetivamente tratado e outro não tratado, apresentando indícios de que a hipótese de suporte comum será atendida. Ademais, é interessante observar como estas variáveis se comportam quando são agregadas por escola, ou seja, observar a

proporção de beneficiados por escola e a proporção estimada de alunos participantes por escola uma vez que este trabalho utilizará regressões por escola, como indicado anteriormente. A tabela 9, assim como o gráfico 2, apresentados a seguir, mostram como ambas as variáveis se relacionam através da correlação e a dispersão das observações, respectivamente.

Tabela 9: Correlação entre a proporção de tratados por escola e a proporção estimada

	Proporção estimada de alunos tratados
Proporção efetiva de tratados	0,77

Gráfico 2: Proporção de aluno tratados por escola x proporção estimada



A tabela 9, assim como o gráfico 2 mostram que, de certa forma, a proporção de alunos beneficiados estimada através do Logit não se distancia muito da real proporção de alunos beneficiados por escola.

4.2. Estimação do efeito marginal de um aumento da proporção de alunos participantes do PBF

A continuidade deste trabalho, assim como antecipado na metodologia, se dedica a estimação do efeito de um aumento da proporção de alunos tratados nas escolas sobre o desempenho médio destas. A tabela 10 apresenta os resultados da estimação da equação (1'')

Tabela 10: Estimativas de OLS para os resultados de interesse

Variável de interesse:			
Taxa de Aprovação	Coeficiente	Rob. Std. Error	
Proporção de tratados	-0.019*	(0.006)	
P(.)	-0.262*	(0.008)	
Const.	96.330*	(0.122)	
Observações	28845	R-squared:	0.098
Taxa de Abandono			
Proporção de tratados	-0.014*	(0.003)	
P(.)	0.113*	(0.004)	
Const.	-0.863*	(0.038)	
Observações	28845	R-squared:	0.085
Nota média em Mat.			
Proporção de tratados	-0.170*	(0.009)	
P(.)	-0.603*	(0.012)	
Const.	203.270*	(0.247)	
Observações	27088	R-squared:	0.235
Nota média em Pt.			
Proporção de tratados	-0.163*	(0.011)	
P(.)	-0.739*	(0.015)	
Const.	227.074*	(0.292)	
Observações	27088	R-squared:	0.222

* p<0,01, ** p<0,05

As estimativas apresentadas na tabela 10 apresentam evidências de que um aumento do número de alunos beneficiados pelo PBF, em média, reduz a taxa de aprovação da escola, assim como a taxa de abandono desta e reduz também as notas médias dos alunos nos exames de proficiência. Portanto, é importante notar

que o coeficiente da variável de interesse é significativo em todos os modelos, mostrando indícios de que a causalidade negativa entre o desempenho e a proporção de alunos tratados em uma escola existe. Os coeficientes apresentados na tabela 10 mostram que, na média, um aumento marginal na proporção de alunos beneficiados pelo PBF (aumento de 1% desta proporção) reduz em 0,17 a média das notas dos exames de proficiência, e também em média, reduz em 0.019 p.p. a taxa de aprovação da escola assim como reduz a taxa de abandono em 0.014 p.p.. Assim, caso a proporção de alunos de uma escola dobre (aumente em 100%), espera-se que sua taxa de aprovação caia 1,9 p.p., o abandono 1,4 p.p. e sua nota em 17 pontos nos exames de proficiência, algo bem semelhante ao que se observa na tabela 5 da seção 2. Neste ponto, é importante ressaltar que os resultados encontrados para as taxas de aprovação e abandono são bastante condizentes o que fora encontrado na literatura até então.

Como mencionado na seção da metodologia, a seguir serão estimados novos modelos contento informações das escolas, além das variáveis anteriormente utilizadas para fins de comparação de resultados. Assim como explicado anteriormente, a comparação destes resultados tem como função identificar se houve viés na estimação dos coeficientes apresentados na tabela 10. A tabela 11 apresenta os resultados das estimações de um modelo simples, onde foi regredida apenas a proporção de tratados contra o resultado de interesse, do modelo descrito pela equação (1'') e, em seqüência, o modelo mais geral apresentado pela equação (5) da seção de metodologia.

Tabela 11: Estimativas de OLS para os resultados de interesse para diferentes modelos

Variável de interesse:			
Taxa de Aprovação	Modelo Simples	(1'')	(5)
Proporção de tratados	-0.143*	-0.019*	-0.020*
	(0.004)	(0.006)	(0.006)
P(.)	-	-0.262*	-0.203*
	-	(0.008)	(0.008)
Const.	92.565*	96.330*	95.274*
	(0.113)	(0.122)	(1.462)
Variáveis da escola	N	N	S
Observações	28845	28845	28840
R-Squared	0.053	0.098	0.126

Taxa de Abandono			
Proporção de tratados	0.040*	-0.014*	-0.012*
	(0.002)	(0.003)	(0.003)
P(.)	-	0.113*	0.086*
	-	(0.004)	(0.004)
Const.	0.756*	-0.863*	-0.084
	(0.049)	(0.038)	(0.517)
Variáveis da escola	N	N	S
Observações	28845	28845	28840
R-Squared	0.028	0.085	0.115
Nota média em Mat.			
Proporção de tratados	-0.467*	-0.170*	-0.178*
	(0.007)	(0.009)	(0.008)
P(.)	-	-0.603*	-0.348*
	-	(0.012)	(0.013)
Const.	194.979*	203.270*	187.975*
	(0.217)	(0.247)	(2.892)
Variáveis da escola	N	N	S
Observações	27088	27088	27083
R-Squared	0.166	0.235	0.359
Nota média em Pt.			
Proporção de tratados	-0.526*	-0.163*	-0.176*
	(0.008)	(0.011)	(0.010)
P(.)	-	-0.739*	-0.415*
	-	(0.015)	(0.016)
Const.	216.920*	227.074*	212.385*
	(0.257)	(0.292)	(3.428)
Variáveis da escola	N	N	S
Observações	27088	27088	27083
R-Squared	0.148	0.222	0.353

* p<0,01, ** p<0,05

Corroborando com os resultados apresentados anteriormente, os coeficientes reportados na tabela 11 novamente apresentam indícios de que um aumento na proporção de alunos participantes do PBF numa escola, em média, reduz o desempenho escolar dos alunos, tanto pela aprovação quanto pela proficiência, além de, também em média, reduzir o abandono. Entretanto, o motivo para que fossem estimados novos modelos, em particular os modelos (5), se calcava na comparação dos coeficientes associados à variável proporção de alunos tratados nas escolas.

Como antecipado neste trabalho, caso o coeficiente relativo à proporção de tratados numa escola divergisse no modelo (1'') do mesmo no modelo (5), resultaria

em uma evidência muito forte de que o coeficiente φ do modelo (1'') sofra de alguma forma de viés. Entretanto, não é isso que se observa ao comparar a magnitude dos coeficientes estimados a partir de (1'') para os estimados por (5) e, isto é, a inclusão das variáveis de escola que, como discutido na metodologia, não devem ser correlacionadas com a participação no programa, alterou somente o coeficiente relacionado à probabilidade média de um aluno na escola participar do PBF, portanto, a exclusão das variáveis da escola nos modelos (1'') não viesaram o coeficiente de interesse. Assim, a alteração somente no coeficiente relativo à $P_s^{\wedge*}(\cdot)$, mostra que a condição $E(b_s^*, \xi_s) = 0$, necessária para que o coeficiente de interesse em (1'') seja consistente, é atendida. Em suma, há indícios de que a probabilidade de participar do tratamento foi bem estimada e, portanto, a participação no PBF independe da escola onde o aluno estuda como desejado.

Mesmo os resultados obtidos na presente seção sendo próximos ao que se esperava previamente a realização deste exercício, a seção seguinte deste trabalho pretende continuar a investigar se ainda há possibilidade de haver algum viés nas estimativas encontradas nesta seção. Para isto, o presente trabalho utilizará alguns testes de robustez de resultados, que serão apresentados detalhadamente a seguir, para entender mais sobre os coeficientes encontrados até então.

5. Teste de Robustez dos resultados estimados

Como citado na seção anterior deste trabalho, esta seção se dedicará a realizar alguns testes de robustez a fim de investigar se os resultados obtidos na seção anterior se sustentam quando realizadas algumas alterações nas estimações do efeito do programa. Para tanto, este trabalho utilizará dois testes: o primeiro utilizando uma base “placebo”, e o segundo utilizando o método de diferenças em diferenças, ambos explicados mais detalhadamente adiante.

5.1. Uso de uma base “placebo”

Considere que em 2001, ano em que o PBF inexistia, o desempenho médio da escola dependia linearmente das características médias de seus alunos, condensadas pela probabilidade $P_s^*(.)$, assim como na equação (1’):

$$Y_s^* = \alpha + \beta \cdot P_s^*(.) + \varepsilon_s \quad (7)$$

onde ε_s representa o termo de erro e, tal como na equação (1’), as características das escolas estão incorporadas em ε_s . A idéia do exercício é incluir a variável “proporção de alunos participantes do PBF em 2008” neste modelo e assim, observar se haviam diferenças pré-existentes entre escolas que viriam a comportar proporções distintas de alunos beneficiários. Caso existam, é possível que estas diferenças tenham se propagado ao longo dos anos causando algum viés na estimação dos efeitos obtidos na seção anterior. Em outras palavras, este exercício pretende investigar se as escolas com mais e menos alunos tratados são realmente comparáveis em 2008. Formalmente, imagine que ao invés de se estimar o modelo (7), que seria o modelo adequado para 2001, fosse estimado um modelo que incluía a proporção que a escola s viria a receber de alunos participantes do PBF, como o mostrado a seguir:

$$Y_s^* = \alpha + \beta \cdot P_s^{**}(\cdot) + \varphi \cdot b_s^* + \varepsilon_s \quad (7')$$

Devido à inexistência do programa em 2001, espera-se que o coeficiente associado à participação no PBF por parte dos alunos, φ , seja estatisticamente insignificante. Caso φ seja diferente de 0 estatisticamente, significa que as escolas que em 2008 receberiam uma proporção maior de alunos beneficiários já tinham desempenho médio distinto daquelas escolas que receberiam proporção menor e, portanto, esta diferença pré-existente ao programa, caso perpetuada ao longo do tempo, pode ter gerado algum viés nas estimativas da seção anterior.

Para a realização deste teste, serão utilizadas duas bases adicionais, o SAEB 2001 e o Censo Escolar 2001. O uso do banco de dados do SAEB se explica pelo fato da Prova Brasil ter criação em 2005, ano em que o PBF já vigorava. O SAEB (Sistema de Avaliação da Educação Básica) na realidade é o exame de proficiência nacional predecessor à Prova Brasil, ambas realizadas pelo INEP, e também avalia proficiência em matemática e português. Como seus resultados possuem a mesma escala que as notas da Prova Brasil, a comparação entre os resultados é direta. A grande distinção entre os exames é que o SAEB utiliza uma amostra de todas as escolas brasileiras, enquanto a Prova Brasil é aplicada em todas as escolas da rede pública com mais de 20 alunos. Em outras palavras, pode-se entender a Prova Brasil como sendo um dos estratos de interesse do SAEB, isto é, o estrato das escolas públicas com mais de 20 alunos para as séries finais dos dois ciclos do EF, entretanto a Prova Brasil é censitária para este estrato. Deste modo, o SAEB continua sendo amostral para as escolas privadas e para as urbanas e rurais com menos de 20 alunos. Isto implica que nem todas as escolas presentes no exercício de 2008 estarão presentes neste presente exercício, apenas as que fazem parte da amostra de escolas públicas do SAEB 2001.

Para realizar este teste para as taxas de aprovação e reprovação utilizou-se o Censo Escolar 2001 para obter estas informações. Entretanto, devido à existência do questionário socioeconômico dos alunos no SAEB 2001, a amostra deste exercício se restringiu às escolas presentes nesta base, uma vez que as informações socioeconômicas do aluno serão necessárias para a atribuição da probabilidade de recebimento do benefício em 2008, como apresentado adiante.

O primeiro passo para a estimação da equação (7') é atribuir a cada aluno da amostra uma probabilidade de receber o benefício do programa, $P_s^{A*}(\cdot)$, mesmo que esse não existisse neste ano, já que esta condensa as características socioeconômicas do aluno. Assim, com base nos coeficientes obtidos na tabela 7 da seção anterior, que apresenta os coeficientes referentes a probabilidade de receber o benefício do PBF, imputou-se para cada aluno presente na amostra do SAEB 2001 uma estimativa de $P_s^{A*}(\cdot)$. Atribui-se também, para cada escola, a proporção observada de alunos beneficiados pelo PBF em 2008. Novamente, uma vez obtidas estas probabilidades, foi feita uma média destas para cada uma das escolas da amostra. A tabela 12 mostra os resultados da estimação de (7') utilizando a base do placebo.

Tabela 12: Estimativas de OLS utilizando o SAEB 2001 e Censo Escolar 2001

Variável de interesse:		
Taxa de Aprovação	Coeficiente	Rob std err.
Proporção de tratados em		
2008	-0.031	(0.040)
P(.)	-0.217*	(0.059)
Const.	93.805*	(2.333)
Observações	636	R-squared: 0.028
Taxa de Abandono		
Proporção de tratados em		
2008	-0.015	(0.023)
P(.)	0.165*	(0.039)
Const.	0.383	(1.657)
Observações	636	R-squared: 0.019
Nota média em Pt.		
Proporção de tratados em		
2008	-0.139*	(0.037)
P(.)	-0.754*	(0.054)
Const.	195.829*	(2.058)
Observações	1240	R-squared: 0.257
Nota média em Mat.		
Proporção de tratados em		
2008	-0.176*	(0.036)
P(.)	-0.686*	(0.053)
Const.	204.357*	(2.105)
Observações	1240	R-squared: 0.239

* p<0,01, ** p<0,05

Os resultados apresentados na tabela 12 mostram que os coeficientes relacionados à participação no PBF são estatisticamente insignificantes no caso da taxa de aprovação e de abandono. Assim, os resultados reportados evidenciam que, antes da existência do PBF, escolas que viriam a receber mais ou menos alunos de famílias beneficiadas pelo programa tinham taxa de aprovação e de abandono semelhantes. Desta forma, a diferença observada destas taxas em 2008 seria realmente proveniente da proporção de alunos beneficiários que viria a estudar nesta escola, corroborando com a evidência encontrada de que há causalidade negativa entre esta proporção e as taxa de aprovação e abandono. Em caminho oposto, os coeficientes relativos à proporção de jovens participantes do programa para os exames de proficiência são estatisticamente significantes. Isto significa que as escolas cuja proporção de alunos tratados é maior em 2008, já padeciam, em média, de algum déficit de proficiência em 2001. Em outras palavras, isto significa que as escolas que em 2008 recebiam mais alunos beneficiados pelo PBF já detinham resultados médios em exames de proficiência em média menor e, portanto, não necessariamente seus resultados piores em 2008 são explicadas pelo fato destas escolas receberem uma proporção maior de alunos beneficiados pelo PBF.

Para efeitos de comparação, estimou-se novamente o modelo (1''), presente na seção anterior, entretanto utilizando somente as escolas que estavam também presentes na amostra do SAEB 2001. Assim, a tabela 13 mostra os resultados para 2008 utilizando a mesma amostra de escolas presente na base de 2001, os resultados das estimações obtidas pelo modelo (7') e os resultados do modelo (7') acrescido das informações sobre as escolas.

Tabela 13: Comparação entre os coeficientes do placebo e da base original

Variável de interesse:				
	Mesmas escolas em 2008	Placebo	Mesmas escolas em 2008 (c/ Variáveis de escola)	Placebo (c/Variáveis de escola)
Taxa de Aprovação				
Proporção de tratados	-0.057*** (0.033)	-0.031 (0.040)	-0.060*** (0.033)	-0.032 (0.039)
P(.)	-0.330*	-0.217*	-0.265*	-0.177*

	(0.041)	(0.059)	(0.045)	(0.065)
Const.	98.581*	93.805*	94.793*	91.746*
	(0.716)	(2.333)	(6.907)	(6.036)
Observações	614	636	614	636
R-Squared	0.131	0.028	0.152	0.033
Taxa de Abandono				
Proporção de tratados	-0.020***	-0.015	-0.016***	-0.012
	(0.010)	(0.023)	(0.010)	(0.022)
P(.)	0.134*	0.165*	0.108*	0.099**
	(0.014)	(0.039)	(0.015)	(0.047)
Const.	-1.066*	0.383	-0.939	7.759
	(0.222)	(1.657)	(1.915)	(4.829)
Observações	614	636	614	636
R-Squared	0.104	0.019	0.132	0.023
Nota média em Pt.				
Proporção de tratados	-0.188*	-0.139*	-0.188*	-0.114*
	(0.037)	(0.037)	(0.036)	(0.043)
P(.)	-0.724*	-0.754*	-0.508*	-0.735*
	(0.057)	(0.054)	(0.061)	(0.086)
Const.	204.981*	195.829*	202.002*	189.774*
	(1.329)	(2.058)	(7.135)	(7.690)
Observações	1164	1240	1164	634
R-Squared	0.247	0.257	0.355	0.267
Nota média em Mat.				
Proporção de tratados	-0.201*	-0.176*	-0.189*	-0.167*
	(0.043)	(0.036)	(0.041)	(0.044)
P(.)	-0.780*	-0.686*	-0.495*	-0.664*
	(0.066)	(0.053)	(0.071)	(0.088)
Const.	225.100*	204.357*	219.475*	198.027*
	(1.533)	(2.105)	(7.601)	(8.807)
Observações	1164	1240	1164	634
R-Squared	0.284	0.239	0.332	0.244

* p<0,01, ** p<0,05, ***p<0,1

Os resultados da tabela 13 mostram que mesmo quando incluídas as variáveis de escola, os resultados obtidos na tabela 12 se sustentam. Novamente, pode-se observar que os coeficientes de interesse para as taxas de aprovação e abandono são estatisticamente insignificantes, fato que, novamente, não ocorre para os parâmetros referentes às notas nos exames de proficiência. Ao observar os coeficientes relacionados à proficiência na tabela 13, é interessante notar que os valores dos coeficientes na base placebo são estatisticamente relevantes em ambas as estimações. Isto indica uma diferença de proficiência pré-criação do PBF entre as

escolas, já que os coeficientes, em magnitude, não se distanciam muito dos parâmetros obtidos na seção anterior. Isto evidencia que a diferença de notas observada em 2008 pode conter algum componente que já determinava notas menores para escolas com maior proporção de alunos tratados. Isto também pode ser visto como não aleatoriedade na “escolha” das escolas por parte dos alunos, isto é, a maioria dos alunos que viriam a receber o benefício do PBF seria alocada em escolas com o desempenho, em média, menor que a média. Apesar de mostrar a possibilidade de viés na estimação do efeito da proporção de alunos tratados sobre as notas médias nos exames, é importante salientar que este teste foi realizado para somente uma pequena parcela das escolas (cerca de 5% das escolas da base original), o que talvez impossibilite a generalização deste resultado para todas as escolas utilizadas em 2008.

5.2. Diferenças em diferenças

Assim como a subseção anterior, esta irá dedicar-se a realização de mais um teste de robustez dos resultados principais deste estudo. A metodologia de diferenças em diferenças consiste basicamente em comparar dois grupos, um tratado e outro de controle, em dois períodos do tempo, um pré-tratamento outro pós. A idéia é que possa haver diferenças anteriores a intervenção, como foi mostrado na subseção anterior, e a comparação dos resultados *ex post* a criação do PBF podem ser “contaminados” por estas diferenças pré-existentes e, portanto, a utilização de um período de tempo anterior à criação do programa pode retirar esta diferença já existente. Isto é, este método é capaz contornar problemas de seleção advindos da não aleatoriedade do tratamento. Assim, se os grupos possuíam trajetórias semelhantes em termos de evolução da variável de interesse, e se os choques que ocorreram durante o período selecionado atingir de maneira uniforme os dois grupos, a comparação da evolução dos resultados do grupo controle com a evolução dos resultados do grupo tratado deve gerar o efeito do tratamento sobre o grupo tratado, tudo mais constante. A seguir a metodologia será apresentada com mais formalidade.

Seja $Y_{s,t}^{D_s}$ o resultado médio para a escola s , onde D_s assume o valor 1 quando a escola tem mais de um quarto de seus alunos beneficiado pelo PBF no ano de 2008 e 0 caso contrário; e t é zero no ano de 2001 e um no ano de 2009. Em outras palavras $Y_{s,0}^0$ denota o resultado médio para a escola s em 2001 para uma escola que em 2008 tem uma proporção de beneficiados inferior a 25% e $Y_{s,1}^0$ o resultado médio para esta mesma escola no ano de 2009. Analogamente, $Y_{s,0}^1$ representa o resultado médio em 2001 para uma escola que tem, em 2008, uma proporção de alunos participantes do PBF maior que 25%, e $Y_{s,1}^1$ o resultado para esta mesma escola em 2009. Desta maneira, novamente assuma que o resultado de interesse médio da escola seja uma função linear, entretanto agora no seguinte formato:

$$Y_{s,t}^{D_s} = \alpha + \gamma \cdot D_s + \rho \cdot D_s \cdot t + \theta \cdot t + \varepsilon_{s,t} \quad (8)$$

Assim, considere agora o resultado médio esperado para uma escola com proporção de beneficiários inferior 25% de seus alunos em 2008, em 2001 é:

$$E(Y_{s,t} / D_s = 0, t = 0) = \alpha$$

O resultado médio para estas mesmas escolas em 2009 será:

$$E(Y_{s,t} / D_s = 0, t = 1) = \alpha + \theta$$

Desta forma, quando se tira a diferença entre estes períodos, obtêm-se o efeito do tempo para as escolas do grupo de controle:

$$E(\Delta Y_s^0 = Y_{s,1}^0 - Y_{s,0}^0 / D_s = 0) = \theta$$

De forma similar, o resultado médio para as escolas com proporção maior que 25% de alunos participantes do PBF em 2008, no ano de 2001, será:

$$E(Y_{s,t} / D_s = 1, t = 0) = \alpha + \gamma$$

E para estas mesmas escolas no ano de 2009:

$$E(Y_{s,t} / D_s = 1, t = 1) = \alpha + \gamma + \rho + \theta$$

Novamente, a diferença entre os períodos para as escolas agora do grupo de tratamento é dada por:

$$E(\Delta Y_s^1 = Y_{s,1}^1 - Y_{s,0}^1 / D_s = 1) = \rho + \theta$$

Agora, repare que a subtração das evoluções dos resultados de ambos os grupos gera o efeito do tratamento para os tratados:

$$E(\Delta Y_s^1 - \Delta Y_s^0) = \rho$$

Note também, que neste tipo de modelagem, é possível incluir variáveis de controle que sejam variantes no tempo, como por exemplo, o índice socioeconômico estimado anteriormente (dado pela probabilidade média dos alunos receberem o benefício do PBF nas escolas) e as variáveis físicas das escolas. Desta forma, a tabela 14 apresenta os resultados da estimação de (8) por POLS:

Tabela 14: Estimativa de Dif-em-dif por POLS

	Taxa de aprovação	Taxa de abandono	Nota média em Pt.	Nota média em Mat.
D*t	-0.199 (1.569)	-0.706 (0.960)	0.655 (1.706)	0.485 (1.914)
D	-1.435 (1.180)	-0.003 (0.722)	-5.312* (1.461)	-6.003* (1.639)
t	-0.476 (2.117)	-1.702 (1.295)	7.437* (2.096)	15.246* (2.351)
P(.)	-0.210* (0.045)	0.097* (0.028)	-0.656* (0.048)	-0.642* (0.053)
Constante	94.213* (5.438)	6.415*** (3.326)	186.160* (6.295)	195.462* (7.062)
Var de escola	S	S	S	S
R2	0.072	0.130	0.491	0.524
Observações	1250	1250	1798	1798

* p<0,01, ** p<0,05, ***p<0,1

A insignificância estatística do parâmetro de interesse ρ em todos os modelos mostrados na tabela 14 mostra evidências de que o “tratamento”, que denota se uma escola tem mais de 25% dos seus alunos participantes do PBF em 2008, não afeta o desempenho escolar médio das escolas. Isto significa que o fato de uma escola ter recebido mais de um quarto de seus alunos beneficiários do PBF, na média, não afetou o desempenho médio desta escola entre 2001 e 2008. Outro ponto interessante destes resultados é o comportamento do coeficiente γ ,

relacionado às diferenças de desempenho pré-tratamento. Note que os coeficientes γ são estatisticamente insignificantes para as taxas de abandono e aprovação, evidenciando que não havia diferenças entre estas taxas para escolas definidas como tratadas e não tratadas em 2001. Entretanto, o contrário ocorre com os coeficientes γ dos modelos para os resultados nos exames de proficiência, que são estatisticamente significantes, o que evidencia que as escolas com maior proporção de tratados em 2008 já detinham notas nos exames de proficiência menores.

Aqui é necessário salientar que o presente modelo utiliza uma variável binária de tratamento (determinada por D_s , que assume o valor um caso mais de um quarto dos alunos da escola sejam beneficiários) e os modelos estimados anteriormente utilizam uma variável contínua para a proporção de alunos tratados. Além disso, a escolha de D_s foi feita de maneira *ad hoc*, isto é, o “tratamento” neste caso foi determinado pela média, sem haver algum motivo teórico para tal escolha. Portanto, a insignificância do parâmetro ρ pode ser proveniente desta divisão em dois grupos de tratamento e, assim como o teste anterior, utiliza somente uma parcela das escolas do exercício principal. Porém, mesmo com tais peculiaridades, os modelos apresentados na tabela 14 novamente apontam para uma diferença de notas em exames de proficiência anteriores a criação do PBF.

Com intuito de reforçar o exercício realizado nesta subseção, considere agora que a variável que determina o tratamento, D_s , seja dada por:

$$D_s = \begin{cases} 0, & \text{se } pbf < 18\% \\ 1, & \text{se } pbf > 32\% \end{cases}$$

Agora, a variável de “tratamento” assume o valor um quando a proporção de alunos participantes do PBF é superior a 32% numa escola e zero caso a escola tenha proporção de alunos beneficiários inferior a 18%. Assim foram retiradas as escolas meio desvio padrão acima e abaixo da média. Desta forma, a tabela 15 apresenta os resultados obtidos para a estimação da equação (8), utilizando a nova definição de grupo de tratamento D_s :

Tabela 15: Estimativa de Dif-em-dif utilizando novo D

	Taxa de aprovação	Taxa de abandono	Nota média em Pt.	Nota média em Mat.
D*t	0.351 (1.993)	-2.310** (1.125)	1.643 (2.180)	0.502 (2.472)

D	-3.100**	1.793**	-7.278*	-7.717*
	(1.560)	(0.880)	(1.948)	(2.209)
t	0.181	-0.518	8.712*	17.493*
	(2.597)	(1.466)	(2.598)	(2.946)
P(.)	-0.181*	0.073**	-0.626*	-0.590*
	(0.052)	(0.029)	(0.057)	(0.064)
Constante	93.716*	6.467***	186.958*	195.388*
	(6.166)	(3.479)	(7.222)	(8.189)
Var de escola	S	S	S	S
R2	0.084	0.147	0.524	0.549
Observações	855	855	1236	1236

* p<0,01, ** p<0,05, ***p<0,1

Assim como na tabela 14, os resultados apresentados na tabela 15 mostram que o efeito do “tratamento” para a taxa de aprovação e para as notas médias em português e em matemática não são estatisticamente significantes. Novamente, estes resultados não refutam os resultados obtidos pelos outros modelos. Entretanto, na estimação utilizando a nova diferenciação de grupos de tratamento e controle, encontrou-se que o parâmetro ρ é estatisticamente significativo e negativo para a equação da taxa de abandono, ao contrário do que ocorreu na tabela 14. Isto significa que, de acordo com o modelo estimado, a participação no tratamento (ter mais de 32% de seus alunos beneficiários do PBF), em média, reduziu a taxa de abandono em -2.310 p.p. Nota-se então que, assim como em todos os resultados obtidos neste trabalho, a proporção de alunos que recebem o auxílio do PBF novamente apresenta uma relação de causalidade negativa com a taxa de abandono das escolas. É interessante observar também nestes modelos que, novamente, há indícios de que havia uma diferença de notas em exames de proficiência entre os grupos em 2001. Além disso, quando comparados aos resultados da tabela 14, percebe-se que estes coeficientes são ainda maiores na tabela 15, isto é, há indícios de que quanto maior a diferença na quantidade de alunos tratados entre as escolas em 2008, maior a disparidade de notas antes da criação do programa.

Em suma, os testes de robustez realizados nesta seção corroboram com os resultados encontrados na seção anterior para as taxas de aprovação e reprovação. Mesmo não confirmando os resultados obtidos para as notas médias nos exames da Prova Brasil da seção anterior, os resultados aqui encontrados também não os refutam, uma vez que o exercício principal deste trabalho é mais completo do que os

testes aqui realizados. Entretanto, os resultados obtidos para os testes de proficiência da presente seção apontam para um fato interessante: escolas que receberam em 2008 uma proporção maior de alunos beneficiados pelo PBF já tinham desempenho médio nos exames de proficiência pior do que aquelas com menor proporção.

6. Conclusão

Este trabalho objetivou encontrar alguma relação de causalidade entre a participação no PBF, aqui agregada por escola e, portanto representada pela proporção de alunos beneficiados por escola, e o desempenho escolar medido pela taxa de aprovação, abandono e pela nota média nos exames de proficiência da Prova Brasil. Para tanto, este se utilizou de uma estimação em dois estágios, onde o primeiro pleiteou estimar a probabilidade de cada aluno presente na amostra participar do PBF, utilizando algumas características socioeconômicas e geográficas, e em um segundo estágio estimar o efeito de um aumento na proporção de alunos tratados em uma escola sobre o desempenho médio dos alunos da mesma. Os resultados provenientes da segunda parte da estimação do efeito do programa sugerem que um aumento da proporção de alunos tratados em uma escola, em média, reduz o desempenho escolar, resultado este que corrobora com alguns outros resultados encontrados na literatura.

Os resultados encontrados neste trabalho mostram indícios de que uma elevação marginal na proporção de alunos beneficiados pelo PBF em uma escola, em média, reduz em 0,019 p.p. a taxa de aprovação da escola assim como reduz a taxa de abandono em 0,014 p.p.. No caso dos exames de proficiência, apesar dos resultados sugerirem que um aumento marginal na proporção de alunos participantes do programa, novamente em média, reduz média das notas em ambos os exames de proficiência em 0,17, os testes de robustez utilizados neste trabalho não endossam estes resultados. Estes testes evidenciam a existência de uma diferença de notas em testes de proficiência pré-criação do PBF entre escolas com diferentes proporções de alunos tratados, isto é, escolas com maior proporção de alunos tratados já padeciam de notas inferiores antes mesmo da criação do PBF. Desta forma, é possível que a magnitude encontrada pelo exercício principal seja superestimada, uma vez que ela incorpora esta diferença pré-existente. Entretanto é necessário reiterar que o exercício principal deste trabalho ainda é mais completo do que os testes de robustez. Estes testes também sugerem que os resultados

encontrados pelo exercício principal para as taxas de aprovação e principalmente para o abandono se sustentam.

Em suma, este trabalho aponta para uma relação causal negativa entre a proporção de alunos tratados em uma escola e as suas taxas de aprovação e abandono, assim como suas notas nos exames de português e matemática da Prova Brasil. Mesmo que os testes de robustez não corroborem com a causalidade negativa entre esta proporção e a proficiência média dos alunos, ainda assim pode-se constatar que as escolas com maior porcentagem de alunos tratados detêm notas piores notas tanto em 2001 quanto em 2008. Assim, fica claro que escolas com maior porcentagem de alunos oriundos de famílias beneficiadas pelo PBF necessitam atenção especial por parte dos gestores de políticas públicas, mesmo que a causalidade negativa entre a proficiência e esta proporção não existisse.

Referências

BEHRMAN, Jere; SENGUPTA, Piyali; TODD, Petra. ***The Impact of PROGRESA on Achievement Test Scores in the First Year.*** International Food Policy Research Institute, PROGRESA Reports, Washington, D.C, 2000.

DAMON, Amy; GLEWWE, Paul. ***Three Proposals to Improve Education in Latin American and the Caribbean: Estimates of the Costs and Benefits of Each Strategy.*** Latin America Development Priorities, Cambridge University Press, pp.45-91, New York - NY, 2009.

FERREIRA, Francisco; LEITE, Phillippe e LITCHFIELD, Julie ***The rise and fall of Brazilian inequality: 1981–2004.*** World Bank Policy Research Working Paper, 3867, Washington, D.C.: World Bank, Washington, D.C.: 2006.

FERRO, Andrea; KASSOUF, Ana Lucia e LEVISON, Deborah. ***The Impact of Conditional Cash Transfer Program in Brazil on Household Work Decisions in Brazil.*** XXXVII Encontro Nacional de Economia - ANPEC 37º, Foz do Iguaçu, 2009.

GLEWWE, Paul e KASSOUF, Ana Lucia. ***O Impacto do Programa Bolsa Família no total de matrículas do ensino fundamental, taxas de abandono e aprovação.*** Seminário Itaú Internacional de Avaliação Econômica de Projetos Sociais, São Paulo, 2010.

de JANVRY, Alain; FINAN, Frederico e SAUDOLET, Elisabeth. ***Local Governance and Efficiency of Conditional Cash Transfer Programs: Bolsa Escola in Brazil.*** Department of Agricultural and Resource Economics, University of California, Berkeley, 2007.

MENEZES FILHO, Naércio. ***Os Determinantes do Desempenho Escolar no Brasil***. Instituto Futuro Brasil. São Paulo: Ibmec, 2007.

PAES DE BARROS, Ricardo; CARVALHO, Mirela; FRANCO, Samuel e MENCONÇA, Rosane. ***Uma Análise das Principais Causas da Queda Recente na Desigualdade de Renda Brasileira***. Texto para Discussão no. 1203. IPEA, Rio de Janeiro, 2006.

PELLEGRINA, Heitor Sandes. ***Impactos de curto prazo do programa bolsa família sobre o abandono e o desempenho escolar do alunado paulista***. Dissertação de mestrado, FEA-USP, São Paulo, 2011.

RESENDE, Anne Caroline; OLIVEIRA, Ana Maria. ***Avaliando Resultados de um Programa de Transferência de Renda: o Impacto do Bolsa-Escola sobre os Gastos das Famílias Brasileiras***. Est. Econ. V.38, n.2, p.235-265, São Paulo, 2008

ROCHA, Sônia. ***Impacto sobre a pobreza dos novos programas federais de transferência de renda***. Revista Economia Contemporânea, v. 9, n. 1, p. 153-185, Rio de Janeiro, 2005.

SOARES, Fabio; SOARES, Serguei; MEDEIROS, Marcelo; OSÓRIO, Rafael. ***Programas de Transferência de Renda no Brasil: Impactos sobre a Desigualdade***. Texto para discussão n° 1228, Brasília: IPEA, 2006.

TAVARES, Priscilla. ***Efeito do Programa Bolsa Família sobre a oferta de trabalho das mães***. XIII Seminário sobre Economia Mineira–Economia, História, Demografia e Políticas Públicas, Diamantina – MG, 2008

TAVARES, Priscilla; PAZELLO, Elaine; FERNANDES, Reynaldo; CAMELO, Rafael. ***Uma avaliação do Programa Bolsa Família: focalização e impacto na***

distribuição de renda e pobreza. Pesquisa e Planejamento Econômico, Rio de Janeiro, v. 39, p. 25-58, 2009