



Caio Cordeiro de Resende

Ensaio em Avaliação de Políticas Públicas

Brasília - DF

Dezembro de 2017

Caio Cordeiro de Resende

Ensaaios em Avaliação de Políticas Públicas

Tese de doutorado submetida ao Departamento de Economia da Universidade de Brasília como requisito para a obtenção do grau de Doutor em Economia.

Universidade de Brasília

Faculdade de Economia, Administração e Contabilidade

Programa de Pós-Graduação em Economia

Orientadora: Ana Carolina Pereira Zoghbi

Brasília - DF

Dezembro de 2017

Nome: Caio Cordeiro de Resende

Título: Ensaio em Avaliação de Políticas Públicas

Tese de doutorado submetida ao Departamento de Economia da Universidade de Brasília como requisito para a obtenção do grau de Doutor em Economia.

BANCA EXAMINADORA

Prof. Dra. Ana Carolina Pereira Zoghbi
Departamento de Economia - UnB
Orientadora

Prof. Dr. Rafael Terra de Menezes
Departamento de Economia - UnB

Prof. Dr. Fernando Boarato Meneguim
Senado Federal e IDP

Prof. Dr. Leonardo Monteiro Monastério
Instituto de Pesquisa Econômica Aplicada e UCB

Aos meus pais, Carlos e Glória,

Ao meu irmão, Igor,

À Manoela.

Agradecimentos

Agradeço à Manoela, minha companheira e primeira leitora, pelo amor, pelo carinho e pela tranquilidade, em horas fáceis e difíceis, e pela paciência e atenção para ler, reler, comentar e criticar diversas versões deste trabalho. Sua companhia e seu apoio tornam simples e fácil todo o resto.

Aos meus pais, Carlos e Glória, pelo amor, pela dedicação e pelo apoio total e incondicional em todas as minhas escolhas. Ao meu irmão, Igor, pela preciosa companhia desde sempre e até sempre.

À minha orientadora, Ana Carolina Zoghbi, pela paciência e dedicação com as quais leu e comentou diversas versões deste trabalho. Suas valiosas críticas, sugestões e correções foram fundamentais para o resultado final desta tese.

Aos membros da banca, Fernando Meneguim, Leonardo Monastério e Rafael Terra pelas críticas e sugestões e, principalmente, pelo estímulo à continuação do trabalho iniciado com esta Tese.

Aos amigos, Luciano, Guilherme, Ivan e Artur, pelas décadas de companhia, amizade e, principalmente, risadas, que tornam a vida mais leve; aos amigos Raphael e Francisco, companheiros inestimáveis na jornada pessoal, profissional e acadêmica; aos amigos Thiago e Ricardo, pelo aprendizado, acadêmico e profissional, nesses diversos anos de convivência; aos amigos de faculdade, Diego, Gustavo, Daniel, Igor(s), Aloisio, Matheus, Fernando, Renata, Carol, Yana, Juliana e Dani, com quem tenho o prazer de compartilhar vitórias e derrotas desde 2003; e aos demais amigos de sempre, que se sabem aqui citados, pelos anos de convivência e aprendizado.

Resumo

Este trabalho consiste em um conjunto de ensaios que avaliam três importantes políticas públicas implementadas recentemente no Brasil. O primeiro ensaio avalia a política nacional de privatização de aeroportos. Utilizamos a metodologia de controle sintético para estimar o contrafactual de como teriam evoluído as receitas dos aeroportos privatizados caso tivessem permanecido sob operação da Infraero. Trata-se de uma aplicação inédita dessa metodologia para avaliação de políticas de privatização. Mostramos que os impactos da privatização nas receitas comerciais são de grande magnitude, estatisticamente significativos e imediatos em todos os aeroportos avaliados. Os resultados mostraram-se robustos a uma série de testes de placebo e falsificações. O segundo ensaio avalia o Programa Mais Educação - PME, principal política de estímulo à educação integral atualmente vigente no Brasil. Neste trabalho, exploramos a descontinuidade existente no critério de elegibilidade do programa a partir de 2012 para avaliar seu impacto em dois indicadores sociais: o trabalho infantil e o desemprego das mães. Além disso, avaliamos o impacto do programa em indicadores relacionados à satisfação dos pais e dos estudantes com a escola, ao comportamento dos estudantes e aos índices de estudantes em recuperação. Não encontramos evidência de impacto significativo do programa em nenhum dos indicadores avaliados. Por fim, o terceiro ensaio analisa os efeitos da inserção de tecnologias de informação e comunicação (TIC) no ambiente escolar no desempenho acadêmico de estudantes. Para isso, realizamos a primeira avaliação de impacto rigorosa do Programa Um Computador por Aluno – Prouca, que forneceu um *laptop* para cada estudante, além de 120 horas de treinamento para docentes, em mais de 300 escolas brasileiras em 2010. Os resultados indicam que, embora o programa tenha impactado positivamente a razão computador por aluno e a utilização de computadores e da internet para fins pedagógicos, não há evidência de efeitos significativos no desempenho acadêmico dos estudantes em testes padronizados de matemática e português e nas taxas de abandono.

Palavras-chave: Avaliação de Impacto; Privatização de Aeroportos; Controle Sintético; Educação Integral; Regressão Descontínua; Tecnologia e Educação; Diferenças-em-Diferenças

Abstract

This study consists of a set of essays that evaluate three important public policies recently implemented in Brazil. The first essay evaluates the national policy of airport privatization. We used a synthetic control approach to estimate the counterfactual of how the revenues would have evolved had the airports remained under Infraero's operation. To the best of our knowledge, this is the first time this approach has been used to evaluate a privatization policy. We have shown that the impacts of privatization on commercial revenues are of great magnitude, statistically significant and immediate at most airports. The results were robust to a series of placebo and falsification tests. The second essay evaluates the effect of a large Brazilian federal program (Mais Educacao) that lengthened the school day from half to full-day shifts on two social indicators: child labor and mother's unemployment. In addition, we evaluated the impact of the program on indicators related to parents' and students' satisfaction with school, student behavior and the rates of students in recovery. We exploit the discontinuity in the eligibility criterion of the program from 2012 to identify the impact of the program. We found no evidence of significant impact of the program on any of the indicators evaluated. Finally, the third essay investigates the effects of information and communication technologies (ICT) in the school environment on student achievement. To this end, we conducted the first rigorous impact evaluation of the One Laptop Per Child – OLPC Brazilian program, which provided a laptop for each student, in addition to 120 hours of training for teachers, in more than 300 schools in 2010. Our results indicate that the program has had a significant impact on the computer-to-student ratio and the use of computers and internet for teaching purposes. Nevertheless, no evidence is found of effects on dropout rates and students' academic performance in Math and Language standardized tests.

Keywords: Impact Evaluation; Airport Privatization; Synthetic Control; Full Time Education; Regression Discontinuity; Technology and Education; Difference in Differences

Lista de Figuras

Figura 1 - Evolução real e contrafactual das receitas comerciais: Aeroportos Reais vs. Aeroportos Sintéticos.....	45
Figura 2 – Impacto estimado da privatização nas receitas comerciais nas 6 especificações	47
Figura 3 – Diferença da evolução das receitas comerciais em Aeroportos Privatizados e Placebos (20 aeroportos).....	50
Figura 4 – Razão do MSPE Pós-privatização e Pré-Privatização (Aeroportos privatizados e controles).....	51
Figura 5 - Evolução real e contrafactual das receitas comerciais: tratamento placebo em 2010	53
Figura 6 – Descontinuidade na Participação no PME em 2012.....	82
Figura 7 - Efeitos do PME no desemprego da mãe	86
Figura 8 – Efeitos do Programa Mais Educação no trabalho infantil.....	88
Figura 9 - Efeitos do PME na Satisfação dos Pais com a Escola.....	91
Figura 10 - Efeitos do PME em Indicadores Comportamentais e na Satisfação dos Estudantes com a Escola	94
Figura 11 - Efeitos do PME em Indicadores Acadêmicos nos Anos Finais	97
Figura 12 – Densidade da Variável de Elegibilidade	99
Figura 13 - Efeitos Placebos do PME (testes em covariadas)	101
Figura 14– Distribuição Nacional do Programa Um Computador por Aluno	134
Figura 15 – Histograma do número de alunos por escola	135

Figura 16 - Impacto do Prouca no Número de Computadores por Escola	152
Figura 17 - Impacto do Prouca no Número de Computadores por Aluno	152
Figura 18 - Impacto do Prouca no Uso do Computador Para Fins Pedagógicos	156
Figura 19 - Impacto do Prouca no Uso da Internet Para Fins Pedagógicos	156
Figura 20 - Impactos do Procura no Desempenho Acadêmico.....	159

Lista de Tabelas

Tabela 1 – Participação das Receitas Comerciais no Total de Receitas no ano imediatamente anterior à privatização (em %)	26
Tabela 2 – Características das Concessões	29
Tabela 3 – Resultados dos Leilões de Aeroportos (em milhões de R\$).....	32
Tabela 4 – Lucro/Prejuízo do Exercício vs. Outorgas Anuais Estimadas (em milhões de R\$)	34
Tabela 5 – Características pré-privatização dos 25 maiores aeroportos operados pela Infraero (média 2004-2012)	37
Tabela 6 – Peso dos Aeroportos do <i>Pool</i> nos Aeroportos Sintéticos.....	42
Tabela 7– Características pré-privatização - média 2004-2012 para o primeiro grupo e 2004-2014 para o segundo grupo (em milhares).....	43
Tabela 8 – Impacto Anual Estimado da Privatização na Evolução das Receitas Comerciais (em milhões de R\$).....	48
Tabela 9 – Comparação entre Modelos.....	54
Tabela 10 – Critérios de Elegibilidade do Programa Mais Educação	68
Tabela 11 – Estimativas dos Efeitos da Elegibilidade para o PME na Participação no Programa em 2012 (primeiro estágio)	80
Tabela 12 – Efeitos do PME no desemprego da mãe (segundo estágio)....	85
Tabela 13 – Efeitos do PME no trabalho infantil (segundo estágio)	88
Tabela 14 – Efeitos do PME na satisfação dos pais com a escola (segundo estágio)	90
Tabela 15 – Efeitos do PME em indicadores comportamentais e na satisfação dos alunos com a escola (segundo estágio).....	93

Tabela 16 – Efeitos do PME em indicadores acadêmicos (segundo estágio)	96
Tabela 17– Efeitos do PME no Emprego das Mães e na Satisfação dos Pais com a Escola (evolução dos indicadores entre 2011 e 2012)	105
Tabela 18– Efeitos do PME em Indicadores Comportamentais, na Satisfação dos Estudantes com a Escola, no Trabalho Infantil e em Indicadores Acadêmicos.....	106
Tabela 19 - Estatísticas Descritivas e testes de médias – Anos Iniciais...	146
Tabela 20 – Estimativas de Impacto do Prouca no número de computadores por escola e por aluno – Anos Iniciais	150
Tabela 21 – Estimativas de Impacto do Prouca no número de computadores por escola e por aluno – Anos Finais.....	151
Tabela 22 – Estimativas de Impacto do Prouca no uso de computadores e internet para fins pedagógicos – Anos Iniciais.....	155
Tabela 23 – Estimativas de Impacto do Prouca no desempenho acadêmico e no abandono – Anos Iniciais e Finais	158
Tabela 24 – Teste de Igualdade de Médias antes do Prouca.....	161

Sumário

APRESENTAÇÃO	14
<hr/>	
<u>1 ENSAIO I: AVALIAÇÃO DE IMPACTO DA POLÍTICA DE PRIVATIZAÇÃO DE AEROPORTOS BRASILEIRA: UMA ABORDAGEM POR CONTROLE SINTÉTICO</u>	<u>22</u>
1.1 INTRODUÇÃO	23
1.2 O PROGRAMA DE PRIVATIZAÇÃO DE AEROPORTOS	25
1.2.1. UMA CARACTERÍSTICA PECULIAR DO SETOR AEROPORTUÁRIO	25
1.2.2. O PROGRAMA DE PRIVATIZAÇÃO DE AEROPORTOS BRASILEIRO: CARACTERÍSTICAS REGULATÓRIAS	27
1.2.3. ÁGIOS EXPRESSIVOS: DIFERENCIAL DE EFICIÊNCIA OU APOSTA NA CAPTURA?	31
1.3 ESTRATÉGIA EMPÍRICA	35
1.3.1. O MÉTODO DE CONTROLE SINTÉTICO PARA ESTUDOS DE CASO	35
1.3.2. BASES DE DADOS E ESTRATÉGIA EMPÍRICA	40
1.4 ANÁLISE DOS RESULTADOS	41
1.4.1. AEROPORTOS SINTÉTICOS: RESULTADOS DAS ESTIMAÇÕES	41
1.4.2. INFERÊNCIA: TESTES DE PLACEBO E FALSIFICAÇÕES	48
1.5 DISCUSSÃO DOS RESULTADOS E CONCLUSÃO	55
1.6 REFERÊNCIAS BIBLIOGRÁFICAS	59
<u>2 ENSAIO II: AMPLIAÇÃO DA JORNADA ESCOLAR E INDICADORES SOCIAIS: UMA AVALIAÇÃO DE IMPACTO DO PROGRAMA MAIS EDUCAÇÃO NO TRABALHO INFANTIL E NO DESEMPREGO DAS MÃES</u>	<u>63</u>
2.1 INTRODUÇÃO	64
2.2 O PROGRAMA MAIS EDUCAÇÃO	65
2.3 BASES DE DADOS	69
2.4 ESTRATÉGIA EMPÍRICA	71
2.5 ANÁLISE DOS RESULTADOS	78
2.5.1. PARTICIPAÇÃO NO PROGRAMA MAIS EDUCAÇÃO	78
2.5.2. PROGRAMA MAIS EDUCAÇÃO E O DESEMPREGO MATERNO	83
2.5.3. PROGRAMA MAIS EDUCAÇÃO E O TRABALHO INFANTIL	87
2.5.4. PROGRAMA MAIS EDUCAÇÃO E A SATISFAÇÃO DOS PAIS COM A ESCOLA	89
2.5.5. PROGRAMA MAIS EDUCAÇÃO E INDICADORES COMPORTAMENTAIS	92
2.5.6. PROGRAMA MAIS EDUCAÇÃO E INDICADORES ACADÊMICOS	95
2.6 TESTES DE ROBUSTEZ	98
2.6.1. TESTE DE DENSIDADE DA VARIÁVEL DE ELEGIBILIDADE	98
2.6.2. TESTES PLACEBO E DIFERENÇA DE MÉDIAS EM COVARIADAS	99
2.7 ESTRATÉGIA EMPÍRICA ALTERNATIVA: UMA ESTIMAÇÃO POR DIFERENÇAS-EM-DIFERENÇAS	100
2.7.1. METODOLOGIA	100
2.7.2. IMPACTO NAS VARIÁVEIS DE INTERESSE	104
2.8 CONCLUSÃO	107

2.9	REFERÊNCIAS BIBLIOGRÁFICAS	109
2.10	APÊNDICE	111
3	<u>ENSAIO III: TECNOLOGIA E DESEMPENHO ESCOLAR: UMA AVALIAÇÃO DE IMPACTO DO PROGRAMA UM COMPUTADOR POR ALUNO</u>	<u>124</u>
3.1	INTRODUÇÃO	125
3.2	REVISÃO DA LITERATURA	127
3.3	O PROGRAMA UM COMPUTADOR POR ALUNO - PROUCA	132
3.4	ESTRATÉGIA EMPÍRICA	137
3.4.1.	MÉTODOS DE PAREAMENTO – <i>PROPENSITY SCORE MATCHING</i> (PSM)	137
3.4.2.	ESCOLHA DAS VARIÁVEIS - O VETOR DE CARACTERÍSTICAS X	140
3.4.3.	DIFERENÇAS-EM-DIFERENÇAS (DD)	141
3.4.4.	A COMBINAÇÃO DOS MÉTODOS: ESCORE DE PROPENSÃO E DIFERENÇAS EM DIFERENÇAS	142
3.5	BASE DE DADOS	143
3.5.1.	FONTES UTILIZADAS	143
3.5.2.	ESTATÍSTICAS DESCRITIVAS	145
3.6	ANÁLISE DOS RESULTADOS	148
3.6.1.	BALANCEAMENTO DO ESCORE DE PROPENSÃO	148
3.6.2.	IMPACTO NO NÚMERO DE COMPUTADORES E DE COMPUTADORES POR ALUNO	148
3.6.3.	IMPACTO NA UTILIZAÇÃO DE COMPUTADORES E INTERNET PARA FINS PEDAGÓGICOS	153
3.6.4.	IMPACTO NO DESEMPENHO ACADÊMICO DOS ESTUDANTES E NAS TAXAS DE ABANDONO	157
3.6.5.	HIPÓTESES PARALELAS E TESTE DE IGUALDADE DE MÉDIAS ANTES DO TRATAMENTO	160
3.6.6.	RESULTADOS HETEROGÊNEOS	162
3.7	CONCLUSÃO	162
3.8	REFERÊNCIAS BIBLIOGRÁFICAS	164
3.9	APÊNDICE	169

Apresentação

Nos últimos anos, observa-se um crescente reconhecimento da necessidade de se aprimorar a implementação de políticas públicas e programas sociais tendo como parâmetro a avaliação de seus resultados. São inúmeras as pesquisas que buscam avaliar políticas e programas com vistas a determinar seus reais efeitos e comparar o custo-benefício de diferentes intervenções. Contudo, isso não é uma tarefa simples. Especialmente porque uma avaliação rigorosa envolve deduzir relações de causalidade que, em virtude da forma como políticas e programas são implementados, podem exigir métodos relativamente sofisticados.

A complexidade técnica dos métodos utilizados para demonstrar causalidade é um obstáculo não apenas para a inserção da prática da avaliação no cotidiano das organizações, mas também para uma comunicação efetiva dos resultados de avaliações para os tomadores de decisões – em geral, políticos e servidores públicos. Particularmente em um trabalho de avaliação de política pública, garantir a comunicação efetiva dos resultados é tarefa tão importante quanto a avaliação *per se*. Por essa razão, buscaremos adotar, nesta apresentação, uma linguagem mais acessível, direcionada não somente a “leitores acadêmicos” e *experts*, mas ao público em geral interessado no estudo de efeitos causais de políticas e programas governamentais.

O principal objetivo desta apresentação é, portanto, oferecer ao leitor não-especializado uma noção, ao menos intuitiva, sobre os motivos que nos levaram a adotar as diferentes técnicas empregadas nos ensaios que compõem este trabalho, além de auxiliá-lo no entendimento dos conceitos, na interpretação dos resultados e na compreensão das limitações metodológicas.

Uma “avaliação de impacto” tem como objetivo avaliar a capacidade de uma política ou um programa alcançarem objetivos específicos. Para isso, são definidos um resultado esperado, um indicador (“variável de interesse”) e um período de tempo em que esse resultado deveria ser alcançado. “Impacto” é nada mais do que a mudança esperada na variável de interesse que pode ser atribuída à política e/ou ao programa. “Tratamento”, por sua vez, é toda política ou programa que se busca avaliar.

A título de exemplo, no Ensaio 2, avaliamos o impacto do Programa Mais Educação - PME (política federal de financiamento à educação integral) - que é, portanto, o nosso tratamento. Já as variáveis de interesse podem ser várias. Diversos trabalhos avaliaram o impacto do PME, por exemplo, no desempenho acadêmico dos estudantes em testes padronizados¹. No Ensaio 2, adotamos como as duas principais variáveis de interesse o trabalho infantil e o desemprego da mãe. Avaliar o PME nesse caso é, portanto, nada mais do que mensurar as mudanças que ocorreram nessas variáveis *por causa* do PME.

Estabelecer a existência de uma correlação – ou uma associação - entre duas variáveis é tarefa relativamente simples. A título de exemplo, analisar a correlação entre a implantação do Mais Educação e a nota dos alunos em testes padronizados ou o trabalho infantil não exige grandes esforços metodológicos. As dificuldades começam quando se deseja extrair uma relação de causalidade entre as duas variáveis, ou seja, quando se busca demonstrar que a nota dos estudantes ou que o percentual de estudantes que trabalham aumentou (ou diminuiu) *por causa* do Mais Educação.

Ao longo dos anos, foram desenvolvidas diversas técnicas estatísticas e econométricas que buscam isolar o efeito causal. Em geral, essas técnicas envolvem a formação de dois grupos:

- Grupo de tratamento: isto é, o grupo de indivíduos e/ou entidades que foram objeto do tratamento – no caso do Ensaio 2, trata-se de grupo de escolas que aderiram ao Programa Mais Educação;
- Grupo de controle: grupo de indivíduos e/ou entidades que não foram objeto da intervenção – no caso do Ensaio 2, trata-se de um subgrupo do grupo de escolas que não aderiram ao Mais Educação.

As dificuldades de realizar uma avaliação bem-sucedida podem ser melhor compreendidas como um problema de *falta de dados*. Isso porque

¹ Vide seção 2.1.

uma avaliação implica deduzir o que teria acontecido caso aquele programa não tivesse sido implementado. No caso específico do PME, avaliar seu impacto no trabalho infantil significa deduzir qual seria o percentual de estudantes que trabalham em determinada escola participante do programa caso a escola não tivesse aderido ao Mais Educação.

Se pudéssemos observar o comportamento da variável de interesse caso o grupo tratado não tivesse participado do tratamento, não haveria dificuldades na avaliação. Bastaria comparar a variável de interesse nos cenários com e sem tratamento e teríamos uma mensuração exata do impacto do tratamento. Como isso não é possível, a construção de um contrafactual (ou seja, de uma base de comparação com relação a qual podemos avaliar o programa) pode ser considerado o principal desafio de uma avaliação de impacto bem-sucedida.

Uma alternativa seria comparar o desempenho das unidades tratadas antes e depois do tratamento. Avaliaríamos, assim, a evolução da variável de interesse ao longo do tempo. Se tomássemos como exemplo as escolas que aderiram ao PME em 2012, isso significaria comparar as variáveis de interesse em 2011 (antes do tratamento) e em 2013 (depois do tratamento). Caso se observe uma diminuição no percentual de estudantes que trabalham em escolas que aderiram ao PME no ano seguinte à adesão, poderíamos deduzir uma relação de causalidade, ou seja, que o PME reduz o percentual de estudantes que trabalham?

A resposta é, claramente, negativa. Essa diminuição no trabalho infantil pode, por exemplo, ter sido generalizada (ou seja, também ter ocorrido em escolas que não aderiram ao PME). Ou ainda pode ser que, no mesmo período, as escolas que aderiam ao PME também tenham sido alvo de outro programa educacional simultaneamente. Como concluir que foi, de fato, o PME o responsável pela diminuição no trabalho infantil, ou seja, que o trabalho infantil diminuiu *por causa* do programa? Nesse caso, podemos ter um problema de “variável omitida”, que ocorre quando a correlação observada entre duas variáveis (implementação do PME e percentual de alunos que trabalham) se deve a uma terceira variável, que não observamos diretamente.

Uma segunda alternativa seria comparar o grupo de tratamento com um grupo controle formado por unidades que não aderiram ao tratamento.

No caso do PME, isso significaria comparar a média de trabalho infantil em escolas do PME com a média em escolas que não participam do programa. Contudo, também nesse caso a comparação não implicaria necessariamente uma relação de causalidade. Isso porque, no momento da seleção para o programa, podem ter sido priorizadas, por exemplo, escolas com maior incidência de trabalho infantil (ou com uma variável correlacionada com o trabalho infantil, como escolas que atendem famílias com situação socioeconômica mais vulnerável). Assim, a relação de causalidade poderia ser inversa: essas escolas foram selecionadas para o PME justamente porque tinham maior trabalho infantil. Trata-se de um caso típico de “causalidade reversa”.

Uma terceira alternativa seria comparar o grupo de tratamento com um grupo de controle que reunia as condições para adesão ao tratamento, mas não o fez. No caso do Mais Educação, isso significaria comparar o desempenho em escolas que aderiram ao PME com aquelas que não aderiram, mas que poderiam tê-lo feito se assim o quisessem - já que a adesão ao programa é voluntária. Nesse caso, contudo, estaríamos diante de um problema conhecido como “auto-seleção”: se as escolas que aderiram possuem, por exemplo, um conjunto de diretores e professores mais motivados e/ou mais qualificados (o que os estimulou a aderir ao programa), a diferença no trabalho infantil pode decorrer dessa motivação/qualificação da direção e do corpo docente da escola - e não do PME.

A existência dos problemas de variável omitida, causalidade reversa e/ou auto-seleção são apenas alguns exemplos das dificuldades de se estabelecer um contrafactual robusto, ou seja, uma base de comparação a partir da qual poderíamos deduzir se o tratamento teve ou não impacto nas variáveis de interesse - ou seja, que nos demonstre o que teria ocorrido com essas variáveis na ausência do tratamento.

Se a seleção das unidades participantes do programa fosse aleatória (por exemplo: sorteio) e um número suficientemente grande de unidades fosse observado, a chamada “lei dos grandes números” garantiria que os dois grupos não seriam sistematicamente diferentes - ou seja, as características dos tratados e dos controles tenderiam ser estatisticamente equivalentes. Nesse caso, seria possível avaliar o impacto do tratamento com relativa facilidade: bastaria comparar as notas médias do grupo de tratamento e de

controle e o impacto do programa seria dado pela diferença entre as notas. Isso porque, com a aleatorização, os dois grupos tenderiam a ser estatisticamente equivalentes em todas as características (observáveis ou não), exceto uma: participarem ou não do tratamento. Avaliações de impacto conduzidas com base em aleatorização são chamadas de “experimentais”.

Contudo, a aleatorização é bastante rara na implementação de políticas públicas, o que dificulta sua avaliação. Trata-se, por exemplo, do caso do Programa Mais Educação, que adotou diversos critérios de priorização, dentre os quais o percentual de estudantes da escola provenientes de família participantes do Programa Bolsa Família. Além disso, a adesão ao programa é voluntária: a escola deve optar por aderir. Nesse caso, a simples comparação entre escolas que implementaram o PME e escolas que não o implementaram tende a gerar uma estimativa enviesada do impacto do programa (já que esses dois grupos serão formados por escolas diferentes entre si).

Essa é a principal razão pela qual lançamos mão de técnicas de avaliação “não-experimentais” para estimar o impacto do Programa Mais Educação, assim como dos dois outros programas avaliados neste trabalho: a Política de Privatização de Aeroportos e o Programa Um Computador por Aluno². Essas técnicas baseiam-se na construção de grupos de controle “artificiais”, a partir de informações provenientes de unidades que não participaram do tratamento. O principal objetivo da construção desses grupos é deduzir o que teria acontecido com as unidades que compõem o grupo de tratamento, caso esse tratamento não tivesse sido implementado.

Em cada um dos ensaios que compõem esta tese, empregamos diferentes metodologias para construção desses grupos de controle. Isso porque as técnicas de avaliação não-experimentais possuem determinadas peculiaridades, que as tornam mais ou menos adequadas a depender do

² Também nesses casos a seleção das unidades tratadas não foi aleatória, como veremos nos Ensaio 1 e 3.

contexto e da forma de implementação de cada política pública e/ou programa governamental.

No primeiro ensaio, utilizamos metodologia conhecida como controle sintético. Por meio dessa técnica, usamos a combinação dos demais aeroportos ainda operados pela Infraero para construir “aeroportos sintéticos”, que se assemelhem aos aeroportos concedidos no período anterior ao da privatização. Nossa expectativa é que esses sintéticos elucidem o que teria ocorrido com os aeroportos privatizados na ausência da privatização. Trata-se de um método adequado para realização de estudos de caso com amostras reduzidas e com apenas uma ou poucas unidades tratadas - o que é justamente o caso da privatização de aeroportos.

No segundo ensaio, empregamos metodologia denominada regressão descontínua. Trata-se de uma técnica cuja aplicação somente é possível quando há uma descontinuidade em um critério de seleção/priorização para participação no programa. Esse é o caso do Programa Mais Educação, que, a partir de 2012, passou a privilegiar a seleção de escolas com mais de 50% dos alunos beneficiários do Programa Bolsa Família (PBF). Essa descontinuidade é usada para avaliar o impacto do programa, comparando a média das variáveis de interesse em escolas tratadas e não-tratadas ao redor dessa descontinuidade (ou seja, escolas com um pouco menos de 50% de alunos beneficiários do PBF e escolas com um pouco mais de 50% de alunos beneficiários do programa).

Finalmente, no terceiro ensaio, adotamos a metodologia do pareamento por escore de propensão com diferenças em diferenças. Para isso, selecionamos um conjunto de características observáveis (ex. número de alunos da escola, localização, nota média dos estudantes, infraestrutura da escola, etc.) e, por meio de um algoritmo, tentamos encontrar, para cada unidade tratada, uma unidade não-tratada que apresente características bastante similares (pareamento). Em seguida, comparamos o grupo tratado e as escolas pareadas do grupo controle antes e depois da implementação do programa (diferenças em diferenças). Esperamos, assim, que o grupo controle pareado, por apresentar características bastante semelhantes ao grupo tratamento, permita-nos deduzir o que teria ocorrido com as unidades tratadas caso o tratamento não tivesse sido implementado.

Nesta apresentação, buscamos introduzir brevemente a temática de avaliação de impacto de políticas e programas governamentais para o leitor não especializado. Nas páginas a seguir, as técnicas empregadas em cada um dos ensaios são explicadas em mais detalhes. Embora essas explicações não possam prescindir de uma linguagem mais técnica, buscamos, sempre que possível, ilustrar os resultados por meio de tabelas e figuras que facilitem sua compreensão.

Ainda que a tarefa de avaliar uma política pública não seja simples, é surpreendente que não haja mais estudos que avaliem, com o rigor necessário, políticas públicas de grande importância para a sociedade brasileira. Particularmente em países em desenvolvimento, onde o custo de oportunidade dos recursos públicos é elevado, a avaliação dos resultados de políticas e programas governamentais deveria ter implicações diretas na sua formulação, implementação e continuidade. Avaliações de impacto rigorosas e bem conduzidas são insumo indispensável para decisões relativas à alocação do orçamento público.

Como vimos, para avaliar os resultados de uma política ou programa específicos, mostrar a existência de correlação entre a política e possíveis resultados é irrelevante; é fundamental demonstrar causalidade. A pergunta fundamental que um tomador de decisão deve – ou deveria – buscar responder é: qual o real impacto dessa política na(s) minha(s) variável(is) de interesse? Ou ainda: o que seria diferente na ausência dessa política e/ou programa?

A análise de políticas públicas baseadas em resultados não tem como prescindir da busca de respostas para questões causais. E é justamente para esse debate que pretendemos contribuir com este trabalho.

Ensaio I

1 ENSAIO I: Avaliação de Impacto da Política de Privatização de Aeroportos Brasileira: Uma Abordagem por Controle Sintético

Resumo

Até 2011, os grandes aeroportos brasileiros eram todos operados por uma única empresa pública: a Infraero. Desde então, o governo federal vem promovendo uma mudança significativa no setor, por meio da privatização dos principais aeroportos do País. Passados seis anos da primeira privatização, não conhecemos qualquer estudo empírico rigoroso sobre seu resultado. Visando preencher essa importante lacuna, neste trabalho avaliamos o impacto da privatização nas receitas comerciais dos aeroportos. Utilizamos a metodologia de controle sintético para estimar o contrafactual de como teriam evoluído essas receitas caso os aeroportos permanecessem sob operação da Infraero. Trata-se de uma aplicação inédita dessa metodologia para avaliação de políticas de privatização. Mostramos que os impactos da privatização nas receitas comerciais são de grande magnitude, estatisticamente significativos e imediatos em praticamente todos os aeroportos privatizados. Os resultados mostraram-se robustos a uma série de testes de placebo e falsificações.

Palavras-chave: Aeroportos; Política de Privatização; Público vs. Privado; Avaliação de Impacto; Controle Sintético.

1.1 Introdução

Nos últimos 6 anos, o mercado aeroportuário brasileiro passou por mudanças significativas, com o avanço do programa de privatizações do Governo Federal. Durante quase 40 anos, a Empresa Brasileira de Infraestrutura Aeroportuária – Infraero foi responsável pela administração dos principais aeroportos do País, no que se poderia caracterizar como um monopólio estatal³. Em apenas três anos, de 2011 a 2014, seis dos maiores aeroportos do País foram concedidos à iniciativa privada, em três rodadas diferentes de concessão. Com isso, cerca de 50% do tráfego aéreo doméstico de passageiros e quase 100% do tráfego internacional são realizados, atualmente, em aeroportos operados pelo setor privado. Este trabalho tem como objetivo realizar a primeira avaliação de impacto da política de privatização de aeroportos brasileira.

Para isso, buscaremos responder à seguinte pergunta: o diferencial de eficiência entre empresas privadas e a Infraero, revelados pelos grandes ágios dos leilões de privatização, de fato, refletem uma eficiência muito superior das empresas privadas na operação dos aeroportos? Se sim, vamos buscar quantificar este diferencial de eficiência. Avaliaremos, assim, se as empresas privadas foram capazes, no curto prazo decorrido desde a privatização, de aumentar significativamente a geração de valor a partir da operação dos aeroportos privatizados.

Para responder a nossa pergunta de pesquisa, avaliaremos a evolução das receitas comerciais dos aeroportos privatizados antes e depois da concessão. Como veremos, no caso dos aeroportos, o diferencial de eficiência entre uma empresa privada e uma empresa pública manifesta-se, basicamente, em duas variáveis: receitas comerciais e custos. Optamos por realizar uma avaliação da primeira variável, visto que, passados poucos anos do início das concessões, os custos dos aeroportos concedidos ainda estão relativamente “inflados” como consequência dos altos investimentos exigidos na primeira fase dos contratos de concessão⁴.

³ A Infraero é a empresa pública federal que, até 2011, era encarregada da administração dos principais aeroportos do País, responsáveis então por 97% dos movimentos do transporte aéreo regular do Brasil (http://www.infraero.gov.br/images/stories/Infraero/Contas/Relatorios/relatorio_anual2011.pdf). Acessado em 10/04/2017.

⁴ Vide Anexo 2 dos contratos. Disponível em: <http://www.anac.gov.br/assuntos/paginas-tematicas/concessoes>. Acessado em 1º de maio de 2017.

Esta análise não é simples. Por um lado, a comparação da evolução das receitas comerciais dos aeroportos privatizados e dos demais aeroportos refletiria não somente o impacto da privatização, mas também o efeito de diferenças entre esses aeroportos nos determinantes dessas receitas. Por outro lado, uma comparação entre as receitas antes e depois da privatização de cada aeroporto não seria capaz de controlar para, por exemplo, mudanças eventuais no contexto econômico pós-privatização que influenciassem a operação dos aeroportos (ex. crise econômica). Em nenhum dos casos teríamos um contrafactual robusto de como teriam evoluído as receitas desses aeroportos caso a privatização não tivesse ocorrido.

Para contornar esse problema, de forma a permitir uma quantificação precisa do diferencial de eficiência das empresas privadas na geração de receitas comerciais dos aeroportos, lançamos mão da metodologia de “controle sintético”. Em termos simples, usamos a combinação dos demais aeroportos ainda operados pela Infraero para construir “aeroportos sintéticos”, que se assemelhem aos aeroportos concedidos no período pré-privatização. Dessa forma, conseguimos gerar um contrafactual robusto de como as receitas comerciais desses aeroportos teriam evoluído caso continuassem a ser operados pela Infraero.

Os resultados do trabalho não deixam dúvidas quanto ao diferencial de eficiência entre as empresas privadas e a Infraero. Chama a atenção não apenas a magnitude desse diferencial, mas também o fato de que apareça, para todos os aeroportos, já no primeiro ano pós-privatização. Tomemos o caso do Aeroporto do Galeão, no Rio de Janeiro. Pelas nossas estimativas, houve um ganho de receitas comerciais de R\$ 174 milhões já no primeiro ano pós-privatização e um ganho acumulado de mais de R\$ 300 milhões nos dois primeiros anos em relação ao que teria ocorrido caso esse aeroporto continuasse sob operação da Infraero. Resultado semelhante foi verificado para os demais aeroportos privatizados. Como veremos, os impactos estimados mostraram-se estatisticamente significativos e robustos a uma série de testes de placebo e falsificações.

Passados mais de 6 anos do primeiro leilão, não conhecemos qualquer estudo empírico rigoroso sobre o impacto das privatizações na operação dos aeroportos brasileiros. Adicionalmente, desconhecemos qualquer aplicação do método de controle sintético para avaliação do impacto de programas de privatização. Nesse sentido, o

presente trabalho inova não somente no sentido de buscar avaliar empiricamente o resultado de uma das mais importantes políticas de privatização em andamento no País, mas também no sentido de ampliar o escopo de aplicação da metodologia de controle sintético.

O restante deste artigo está estruturado em 4 seções. Na seção 1.2, apresentaremos breve panorama do programa de privatização aeroportuária no País e de sua implementação. A seção 1.3 discutirá, com mais detalhes, a estratégia empírica, o modelo econométrico e as bases de dados utilizadas. Na seção 1.4, analisaremos os principais resultados, bem como sua robustez por meio de uma série de testes de placebo. A seção 1.5 discutirá os resultados alcançados, bem como concluirá o trabalho.

1.2 O Programa de Privatização de Aeroportos

1.2.1. Uma Característica Peculiar do Setor Aeroportuário

As receitas de um aeroporto podem ser classificadas em duas categorias: aeronáuticas e não-aeronáuticas (ou comerciais). As primeiras remuneram a prestação do serviço público *strictu sensu*, ou seja, o processamento de passageiros, de voos e de cargas. No Brasil, conforme previsto na Lei nº 6.009, de 1973, existem seis tipos de tarifas aeroportuárias: embarque, conexão, pouso, permanência, armazenagem e capatazia. Em tese, as duas primeiras remuneram a utilização da infraestrutura pelos passageiros, a terceira e a quarta pelas aeronaves e as duas últimas a infraestrutura para operação de cargas⁵. Já as receitas comerciais abarcam uma série de outros serviços oferecidos pelos aeroportos: publicidade, aluguéis de espaço, estacionamento, combustível para aeronaves, *duty free*, alimentação, hotéis, entre outros.

Nesse aspecto, o setor difere bastante de outros setores de infraestrutura, como, por exemplo, rodovias. Neste, as receitas do concessionário advêm basicamente do pedágio, que reflete a demanda dos motoristas pela utilização da infraestrutura (variável que se encontra, em grande medida, fora do controle do concessionário). Assim, o diferencial de eficiência entre uma empresa pública e uma empresa privada no setor de rodovias tende a refletir, basicamente, sua estrutura de custos. No caso do setor

⁵ Dizemos em tese porque a tarifa de conexão é cobrada da empresa e não do passageiro.

aeroportuário, essa diferencial reflete, ademais, a capacidade de geração de receitas comerciais da empresa, o que poderia justificar a existência de ágios expressivos no leilão⁶.

Interessante notar que, atualmente, em grande parte dos aeroportos do mundo, as receitas comerciais são superiores às receitas aeronáuticas (ACI, 2008; ATRS, 2011)⁷. Esse, contudo, não é um fenômeno recente. Já no início da década de 1990, segundo Doganis (1992), as receitas comerciais já representavam 75%-80% das receitas totais de aeroportos médios e grandes nos Estados Unidos. Cenário semelhante se verificava na *British Airports Authority* - BAA no início da década de 1990⁸: as receitas comerciais representavam cerca de 60% das receitas totais dos três aeroportos de Londres (Heathrow, Gatwick e Stansted). Adicionalmente, diversos estudos demonstram que operações comerciais tendem a ser mais lucrativas que operações aeronáuticas (Jones et al., 1993; Starkie, 2001; Forsyth, 2003)⁹.

Esse, contudo, não era o caso dos aeroportos brasileiros no período pré-privatização. Como se nota na Tabela 1, em nenhum dos aeroportos privatizados, as receitas comerciais eram superiores a 50% das receitas totais. Na média, as receitas comerciais respondiam por, apenas, 33,5% das receitas totais no ano imediatamente anterior à privatização.

Tabela 1 – Participação das Receitas Comerciais no Total de Receitas no ano imediatamente anterior à privatização (em %)

Aeroporto	% de Receitas Comerciais
BSB	44%
GRU	38%
VCP ¹⁰	12%
ASGA	34%

⁶ A presença de receitas comerciais desreguladas é uma peculiaridade do setor aeroportuário, com desdobramentos em várias áreas. O impacto da presença de receitas comerciais para a competição intermodal e para os modelos regulatórios do setor foram discutidos, em detalhes, em Resende, Caldeira, & Fonseca (2016).

⁷ Para uma discussão sobre a importância das receitas comerciais para aeroportos, vide Graham (2009).

⁸ A BAA foi a primeira operadora de aeroportos pública privatizada no mundo, em 1987.

⁹ Segundo Jones et al. (1993), a margem operacional das tarifas aeronáuticas atingiu 7% para os três aeroportos da BAA. A margem operacional das receitas comerciais, por sua vez, atingiu 64%.

¹⁰ O aeroporto de Viracopos tem a peculiaridade de ser o segundo maior aeroporto de cargas do País, o que explica a baixa representatividade das receitas comerciais nas receitas totais.

CNF	38%
GIG	35%

Fonte: Elaboração própria com base nos dados da Infraero.

Há indícios, assim, de que as receitas comerciais eram subexploradas, o que pode explicar a existência de um grande diferencial de eficiência entre empresas públicas e privadas na operação desses aeroportos. Na seção 1.4 deste trabalho, buscaremos avaliar se esse diferencial, de fato, existe. Antes, contudo, veremos algumas características importantes do programa de privatização de aeroportos no Brasil.

1.2.2. O Programa de Privatização de Aeroportos Brasileiro: Características Regulatórias

O programa de privatização de aeroportos brasileiro teve início em 2011, com a privatização do Aeroporto Internacional de São Gonçalo do Amarante - ASGA ¹¹, distante 40km de Natal, no Rio Grande do Norte. A concessão do ASGA destacou-se por se tratar de um projeto *greenfield* – ou seja, um novo aeroporto deveria ser inteiramente construído pela empresa vencedora do leilão¹². Assim, apesar de o leilão ter sido realizado em agosto de 2011, o aeroporto somente entrou, de fato, em operação em meados de 2014. Para fins deste trabalho, consideramos 2015 com o primeiro ano pós-privatização.

A segunda rodada de concessões foi realizada em fevereiro de 2012 e incluiu o Aeroporto Internacional de Brasília/DF, o Aeroporto Internacional de Viracopos/SP e o Aeroporto Internacional de Guarulhos/SP. É importante notar que as Concessionárias não assumem o aeroporto imediatamente após o leilão. A concessão, de fato, teve início somente em julho de 2012, momento no qual se iniciou um período de transição de seis meses, no qual a administração dos aeroportos foi realizada conjuntamente com a Infraero. Assim, para efeitos deste trabalho, os aeroportos somente passaram a ser integralmente administrados pelas Concessionárias privadas somente em 2013¹³.

¹¹ Houve, anteriormente, casos pouco conhecidos de transferência de pequenos aeroportos estaduais e municipais à iniciativa privada, como o aeroporto de Cabo Frio, no Rio de Janeiro, o de Fernando de Noronha, em Pernambuco, o de Porto Seguro, na Bahia, e o de Bonito, em Mato Grosso do Sul.

¹² No caso do ASGA, a pista de pouso/decolagem foi construída pelo poder público.

¹³ É importante ressaltar que isso pode gerar um viés de baixa na estimativa do impacto da privatização na evolução das receitas comerciais. Isso porque adotaremos o ano de 2012 como o último ano pré-

Em novembro de 2013, foi realizada a terceira rodada, na qual foram concedidos os aeroportos internacionais de Confins, em Belo Horizonte, e do Galeão, no Rio de Janeiro. O modelo foi semelhante ao da segunda rodada. Assim, a concessão teve início formal somente em maio de 2014, seguido por um período de 6 meses de transição das operações. Adotando raciocínio semelhante ao das primeiras rodadas, consideraremos o ano de 2015 como o primeiro no qual o aeroporto esteve sob administração do concessionário privado¹⁴.

Os leilões dos aeroportos foram realizados em duas etapas. Na primeira, os proponentes deveriam apresentar, em volumes lacrados, suas propostas econômicas para cada aeroporto. As propostas foram, então, ordenadas pelo critério de maior Valor da Contribuição Fixa (montante anual pago pela concessionária ao governo pelo direito de explorar o aeroporto). Em seguida, teve início o leilão em viva-voz, do qual participaram as empresas cujas ofertas estavam entre as três maiores para cada Aeroporto e as empresas cujo valor da oferta fosse igual ou superior a 90% do valor da maior oferta.

Em alguns casos, como Galeão, não houve sequer lances no leilão viva-voz. Isso porque a oferta do consórcio vencedor (R\$ 19 bilhões) foi mais de 30% superior à oferta do segundo colocado (R\$ 14,5 bilhões), que não teve interesse em participar do viva-voz. O mesmo ocorreu no aeroporto de Guarulhos, no qual, já na primeira etapa, a oferta vencedora (R\$ 16 bilhões) foi mais de 25% superior à segunda colocada (R\$ 12,9 bilhões).

Os valores da Contribuição Fixa devem ser pagos em parcelas anuais, reajustadas pela inflação. Deve-se notar que o prazo da concessão varia entre os aeroportos, o que torna os valores totais de outorga não comparáveis entre si. De forma a permitir a comparação, incluímos na Tabela 2 o valor de outorga fixa anual estimado.

Além da Contribuição Fixa, o contrato prevê o pagamento da chamada “Contribuição Variável”, que consiste em um percentual da receita bruta. Esse percentual também varia de acordo com o aeroporto, conforme pode ser observado na Tabela 2¹⁵.

privatização. Contudo, nos 6 meses de operação conjunta transcorridos entre julho e dezembro, as Concessionárias privadas já podem, em tese, ter atuado no sentido de incrementar as receitas comerciais dos aeroportos concedidos.

¹⁴ Em 2017, foi realizada nova rodada de concessão, que incluiu quatro aeroportos: Florianópolis, Fortaleza, Porto Alegre e Salvador. A análise dessas concessões, contudo, foge ao escopo desse trabalho.

¹⁵ Esse percentual pode, ainda, subir em determinados cenários. Para mais detalhes, vide seção intitulada “Da Contribuição ao Sistema” dos Contratos de Concessão. Disponível em:

Essa contribuição torna ainda mais significativa a outorga anual a ser paga pelas Concessionárias, em relação os valores previstos na coluna Outorga Fixa Anual. A título ilustrativo, em 2014, somadas outorga fixa e variável, a contribuição total do Aeroporto de Guarulhos alcançou cerca de R\$ 1,1 bilhão, a de Brasília, R\$ 220 milhões e a de Viracopos, R\$ 170 milhões.

Tabela 2 – Características das Concessões

Aeroporto	Prazo da Concessão	Outorga Fixa Anual (em milhões de R\$)	Outorga Variável Anual (Estimada)
BSB	25 anos	180	2%
GRU	20 anos	811	10%
VCP	30 anos	127	5%
CNF	30 anos	60	5%
GIG	25 anos	760	5%
ASGA	25 anos	10	-

Fonte: Elaboração própria com base nos dados da Anac.

Sob o ponto de vista regulatório, o governo optou por um modelo de regulação por tarifa-teto (*price cap*). Nesse modelo, as tarifas aeroportuárias são reajustadas anualmente pela seguinte fórmula:

$$P_t = P_{t-1} * (1 + IPCA_t) * (1 - X_t) * (1 - Q_t) \quad (1)$$

em que P_t e $IPCA_t$ são, respectivamente, o preço da tarifa e o índice de inflação¹⁶ no período t ; X_t é o chamado “fator de produtividade”, que é igual ou superior a zero nos cinco primeiros anos de contrato e tem como objetivo compartilhar os ganhos de

<http://www.anac.gov.br/assuntos/paginas-tematicas/concessoes/guarulhos/arquivos/contrato-de-concessao>. Acessado em 5 de julho de 2017

¹⁶ No caso, é adotado como referência o Índice Nacional de Preços ao Consumidor Amplo – IPCA.

produtividade e eficiência do aeroporto com o usuário; e Q_t é o “fator de qualidade do serviço”, obtido mediante avaliação o cumprimento dos Indicadores de Qualidade de Serviço estabelecidos nos Contratos (varia entre 7,5% e -2%, conforme o cumprimento das metas)¹⁷.

Como as metas de qualidade do serviço são bastante altas e a bonificação por atingir todas as metas chega a, no máximo, 2%, a tendência é que as receitas tarifárias sejam reajustadas anualmente por valores inferiores à inflação, o que denota um limite claro para o crescimento das receitas tarifárias dos aeroportos concedidos. Essa é uma característica importante da privatização desses aeroportos, uma vez que o ganho estimado nas receitas comerciais ocorre em um contexto em que o custo de utilização dos aeroportos pelos passageiros não aumenta em termos reais.

Os modelos de regulação de aeroportos podem ser classificados, ainda, em dois tipos: *single till* ou *dual till*. No modelo mais comum, o *single till*, as receitas comerciais são também reguladas, existindo, assim, um subsídio cruzado entre as atividades comerciais e as atividades aeronáuticas (o lucro proveniente de atividades comerciais é considerado pelo regulador na definição das tarifas aeroportuárias, o que diminui o incentivo do regulado para geração de receitas comerciais). Já no *dual till*, as receitas comerciais são desreguladas, ou seja, não são consideradas no cálculo das tarifas. Esse é o caso em que o diferencial de eficiência entre empresas públicas e privadas também se manifesta nas receitas comerciais. Isso porque o agente privado tem grandes incentivos para investir na geração de receitas comerciais, uma vez que não são reguladas pelo agente público. Trata-se do modelo de regulação adotado no Brasil¹⁸.

O contrato inclui, ainda, a previsão de investimentos obrigatórios, tanto de curto prazo, a serem realizados imediatamente após o início da operação, quanto de médio e longo prazo, regulado por meio de gatilhos de investimentos, definidos de acordo com a

¹⁷ Para detalhes, vide, por exemplo, o Capítulo VI e os Anexos 4 e 11 dos Contratos de Concessão de Brasília, Guarulhos e Viracopos. Disponível em: <http://www.anac.gov.br/assuntos/paginas-tematicas/concessoes/guarulhos/arquivos/contrato-de-concessao>. Acessado em 5 de julho de 2017.

¹⁸ Ainda que o modelo *single till* seja atualmente predominante, uma série de estudos produzidos tanto por acadêmicos (Forsyth, 2002; D. Starkie & Yarrow, 2000; Oum, Zhang, & Zhang, 2004; David Starkie, 2008) quanto por autoridades reguladoras (CAA, 2000; Productivity Commission, 2011) têm defendido o modelo de regulação *dual till*, principalmente no caso de aeroportos congestionados. A discussão sobre modelos de regulação de aeroportos é realizada em detalhes em Resende, Caldeira, & Fonseca (2016).

movimentação de passageiros e com os Parâmetros Mínimos de Dimensionamento definidos no contrato¹⁹.

Uma última característica importante dos contratos de concessão é a previsão de que a Infraero seja sócia minoritária de todos os consórcios vencedores, com 49% de participação (com exceção do Aeroporto de Natal). Essa é outra característica da privatização que pode levar a um viés de baixa nas nossas estimativas. Isso porque pode ter havido um “efeito aprendido” na Infraero, ao absorver certas práticas das Concessionárias privadas e levá-las aos demais aeroportos da rede. Assim, a privatização pode ter afetado a evolução das receitas comerciais de aeroportos não-privatizados, levando-nos a superestimar o cenário contrafactual, ou seja, o que teria acontecido caso a privatização não tivesse ocorrido.

1.2.3. Ágios Expressivos: Diferencial de Eficiência ou Aposta na Captura?

Uma característica marcante das concessões dos aeroportos no Brasil foram os ágios expressivos observados nos leilões, como podemos observar na Tabela 3.

Como se nota, na primeira rodada de concessões, na qual foi concedido o Aeroporto de São Gonçalo do Amarante - ASGA (que atende a região de Natal, Rio Grande do Norte), o ágio atingiu 229%. As ofertas vencedoras da segunda rodada de concessões, que incluiu os aeroportos de Guarulhos - GRU, Viracopos -VCP e Brasília - BSB, somaram cerca de R\$ 24,5 bilhões, o que representou um ágio médio de 347%, considerado o valor mínimo de R\$ 5,5 bilhões exigido pelo governo. Por fim, na terceira rodada, a arrecadação total alcançou R\$ 20,8 bilhões, o que significou um ágio de 252% em relação ao valor mínimo exigido pelo governo de R\$ 5,9 bilhões.

¹⁹ Vide Anexo 2 (Plano de Exploração Aeroportuária -PEA) e Anexo 4 (Tarifas) do Contrato de Concessão dos aeroportos de Brasília, Guarulhos e Viracopos. Disponível em: <http://www2.anac.gov.br/gru-vcp-bsb/>. Acessado em 15/08/2015.

Tabela 3 – Resultados dos Leilões de Aeroportos (em milhões de R\$)

Aeroporto	Valor mínimo	Oferta Vencedora	Ágio	Consórcio Vencedor	Principais Empresas do Consórcio
ASGA	52	170	229%	Inframérica	Infravix Participações e Corporación America
BSB	580	4.500	674%	Inframérica	Infravix Participações e Corporación America
GRU	3.400	16.200	374%	Invepar	Invepar Participações e Airport Company South Africa
VCP	1.500	3.800	160%	Aeroportos Brasil	Triunfo Participações, UTC Participações e Egis Airport Operacional
CNF	1.100	1.820	66%	AeroBrasil	CCR e a operadora suíça Flughafen Zurich AG
GIG	4.800	19.000	294%	Aeroportos do Futuro	Odebrech Transport e Changi Airport

Fonte: Elaboração própria com base nos dados da Anac.

Os ágios expressivos alcançados em praticamente todos os aeroportos privatizados podem ser interpretados de duas formas. A “hipótese positiva” é que esses ágios demonstrem simplesmente que o mecanismo de leilão está atingindo seus resultados: a partir de uma estimativa conservadora do valor dos aeroportos realizada pelo governo, o leilão está auxiliando na seleção da empresa mais eficiente (ou que se acredita mais eficiente). Nessa hipótese, os ágios refletiriam o diferencial de eficiência entre as Concessionárias privadas em relação à Infraero, cujos dados históricos foram utilizados como base para as estimativas de valor mínimo do Governo Federal.

A “hipótese negativa” é que o leilão esteja auxiliando não na seleção da empresa mais eficiente, mas sim naquela que acredita ter maiores chances de renegociar futuramente o contrato no caso de os valores propostos mostrarem-se inviáveis. Em outras palavras, os ágios expressivos seriam um indicativo de que as empresas estão apostando numa captura futura do agente público, o que facilitaria uma revisão futura do contrato de concessão, caso os termos se mostrem desvantajosos para a empresa. As ofertas elevadas seriam, assim, resultado de uma espécie de “corrida a qualquer custo” pelos aeroportos, uma vez que, mesmo que as concessões não se mostrem lucrativas, acreditar-se-ia na possibilidade de revisão dos termos acordados. Nesse cenário, quanto maior a

probabilidade que se confere no êxito de uma renegociação futura dos contratos, menor a importância que se confere ao valor oferecido no leilão.

A probabilidade de ocorrência de uma ou outra hipótese certamente varia com o contexto institucional do País. Em um cenário caracterizado por instituições governamentais fortes e confiáveis e baixa corrupção, a aposta na hipótese negativa mostra-se uma opção mais arriscada, diminuindo as chances de sua ocorrência. Por outro lado, em um cenário caracterizado por instituições governamentais relativamente fracas e alta corrupção, aumenta a probabilidade de que os agentes engajem em um comportamento de risco, apostando na capacidade de captura do agente público.

Escândalos de corrupção recentes revelados pela Operação Lava Jato levam-nos a crer que o Brasil pode estar mais próximo do segundo caso. Chama a atenção, nesse sentido, o fato de os grandes protagonistas dos escândalos revelados pela operação serem justamente os principais sócios dos Consórcios ganhadores dos aeroportos privatizados, ou seja, as grandes construtoras nacionais (vide Tabela 3). Assim, não se pode descartar a hipótese de que o mecanismo de leilão tenha funcionado não como um instrumento para atração da firma mais eficiente, mas sim para atração da firma mais propensa ao risco - no caso, mais disposta a apostar na chance de revisão futura dos termos do contrato.

Outro indício relevante é fato, de menos de dois anos após a realização do leilão, a Concessionária do Aeroporto do Galeão, liderado pela empresa Odebrecht, já ter solicitado revisão do Contrato de Concessão. Esse pedido foi acompanhado por solicitações semelhantes das demais Concessionárias. O Governo Federal chegou a cogitar a revisão dos contratos, contudo, posteriormente, desistiu da ideia, ao menos temporariamente ²⁰.

No cenário que consideramos positivo, os ágios do leilão podem ser considerados bons indícios de que não mais eficiente é uma empresa privada em relação à empresa pública. Na Tabela 4, comparamos o lucro/prejuízo dos aeroportos concedidos no ano imediatamente anterior à concessão e na média dos últimos três anos com a outorga fixa anual estimada de cada um desses aeroportos.

²⁰ Vide <http://exame.abril.com.br/negocios/concessionarias-pedem-r-2-5-bi-por-revisao-de-contrato/> ou <http://odia.ig.com.br/economia/2017-04-14/governo-pretende-aprovar-revisao-de-outorgas-do-galeao.html> . Acessado em 20/04/2017.

Tabela 4 – Lucro/Prejuízo do Exercício vs. Outorgas Anuais Estimadas (em milhões de R\$)

Aeroporto	Ano/Período de Referência	Lucro/Prejuízo Líquido	Outorga Fixa Anual (Estimada)
BSB	2011	50.3	180
	2009-2011	27.9	
GRU	2011	425.2	811
	2009-2011	361.2	
VCP	2011	97.0	127
	2009-2011	70.4	
CNF	2013	-83.2	60
	2011-2013	0.5	
GIG	2013	-110.1	760
	2011-2013	12.4	
ASGA	2010	-2.5	10
	2008-2010	-7.1	

Fonte: Elaboração própria com base nos dados da Infraero e da Anac.

Assim, enquanto a Infraero lucrou, em média, R\$ 361 milhões com o Aeroporto de Guarulhos nos três anos anteriores à concessão, o Concessionário privado comprometeu-se a pagar ao Governo Federal, anualmente, somente à título de outorga fixa pela concessão, cerca de R\$ 800 milhões, mais de duas vezes o lucro médio da Infraero nos três anos anteriores à Concessão²¹. No caso de Brasília, os valores são, proporcionalmente, ainda mais expressivos. Enquanto o lucro médio da Infraero nos três anos anteriores foi de R\$ 28 milhões, a Concessionária se comprometeu a pagar anualmente quase R\$ 200 milhões ao governo a título de outorga – o que corresponde a cerca de 7 vezes o lucro médio da Infraero. Cenário semelhante se desenha em praticamente todos os aeroportos concedidos, como se nota na Tabela 4.

Os valores expressivos de outorga chamam ainda mais atenção quando se considera que as Concessionárias i) deverão implementar um programa de investimentos

²¹ Como vimos na seção 1.2.2, os contratos preveem, além da outorga fixa, o pagamento de uma outorga variável, calculada em função da receita bruta dos aeroportos – razão pela qual, em 2014, o Concessionário de Guarulhos pagou cerca de R\$ 1,1 bilhão a título de outorga.

obrigatórios muito mais audacioso que o da Infraero²²; e ii) devem recolher aos cofres públicos diversos impostos que a Infraero, na condição de empresa pública, tinha imunidade, tal como o Imposto sobre serviços de qualquer natureza (ISS) e Imposto de Renda de Pessoa Jurídica (IRPJ).

Interessante notar, ainda, que os lances mínimos do leilão são definidos pelo governo com base em Estudos de Viabilidade Técnica, Econômica e Ambiental – EVTEAs. Nesses estudos, o governo realiza seu próprio *valuation* dos aeroportos, de forma a definir um valor mínimo aceitável por cada um deles. Ainda que baseados no histórico de operação da Infraero, os EVTEAs incluem um diferencial de eficiência que considera o fato de o aeroporto passará a ser administrado por um ente privado. Logo, os ágios verificados nos leilões são ainda mais significativos, já que têm como base de comparação estudos que, ao menos em teoria, já levariam em consideração um diferencial de eficiência entre a empresa pública e as empresas privadas.

1.3 Estratégia Empírica

1.3.1. O método de controle sintético para estudos de caso

Para avaliar o impacto das privatizações nas receitas comerciais dos aeroportos utilizaremos o método de controle sintético, introduzido por Abadie & Gardeazabal, (2003), com desenvolvimentos posteriores em Abadie, Diamond, & Hainmueller (2010) e Abadie, Diamond, & Hainmueller (2014). Trata-se de um método relativamente novo e que, segundo Athey & Imbens (2016), representa a principal inovação na literatura de avaliação de impacto nos últimos 15 anos.

Uma das principais aplicações do método é, justamente, a realização de estudos de caso com amostras reduzidas e com apenas uma ou poucas unidades tratadas. Trata-se, exatamente, do caso da privatização dos aeroportos, no qual temos apenas 6 unidades tratadas e 19 unidades no grupo de controle²³.

²² A título de exemplo, os investimentos estimados em Guarulhos nos 20 anos da concessão são de R\$ 4,6 bilhões, enquanto a Infraero investiu, nos 8 anos anteriores à concessão, cerca de R\$ 250 milhões.

²³ Como veremos a seguir, consideramos somente os dados relativos aos 25 maiores aeroportos do País operados pela Infraero em 2011.

A intuição por trás do método é construir, para cada aeroporto privatizado, um “aeroporto sintético”, composto por uma média ponderada de aeroportos não-privatizados. A evolução das receitas comerciais desses aeroportos sintéticos, caso o método seja bem aplicado, demonstra como as receitas comerciais dos aeroportos privatizados teriam evoluído, caso a privatização não tivesse ocorrido.

Uma das grandes vantagens do controle sintético é que a escolha das unidades que comporão o aeroporto sintético é realizada de forma objetiva, com base nos dados disponíveis (*data driven*), o que reduz a discricionariedade do pesquisador na escolha das unidades de comparação. Adicionalmente, como, na prática, é muito difícil encontrar uma única unidade não tratada (ex. um aeroporto não privatizado) cujas características se aproximem das características mais importantes da unidade tratada (ex. aeroporto privatizado), o método gera uma combinação ponderada de unidades, que, em geral, forma uma base de comparação melhor do que qualquer uma das unidades não tratadas, consideradas individualmente.

Construído o aeroporto sintético dessa forma e desde que as diferenças entre o aeroporto real e o aeroporto sintético antes do tratamento sejam mínimas, o efeito estimado da privatização será dado pela diferença entre as receitas comerciais do aeroporto privatizado e as do aeroporto sintético no período pós-privatização.

A seguir, apresentaremos o método de controle sintético de forma mais rigorosa. Os conceitos e a notação desta seção têm por base Abadie & Gardeazabal (2003) e Abadie, Diamond, & Hainmueller (2010).

Seja $J+1$ os 20 maiores aeroportos do País operados pela Infraero em 2011. Suponha que somente o primeiro aeroporto seja afetado pela privatização. Nesse caso, o restante dos aeroportos (não privatizados) formam o grupo de controle – que chamaremos de “*pool* de doadores”.

Vamos definir como T o número de anos no qual observamos os aeroportos e T_0 como o último período antes da privatização, de forma que $1 < T_0 < T$. Vamos definir como Y_{it} o valor das receitas comerciais do aeroporto i no período t , Y_{it}^N o valor das receitas comerciais caso o aeroporto não tenha sido privatizado e Y_{it}^P o valor caso o aeroporto tenha sido privatizado.

O impacto da privatização do aeroporto i no período t será dado por

$$\alpha_{it} = Y_{it}^P - Y_{it}^N \quad (2)$$

Vamos definir, ainda, D_{it} como uma variável *dummy* igual a 1 se o aeroporto foi privatizado e igual a 0, caso contrário. Então podemos reescrever (2) como:

$$Y_{it}^P = Y_{it}^N + D_{it}\alpha_{it} \quad (3)$$

À primeira vista, o cálculo do impacto da privatização parece simples. Contudo, note que, para os aeroportos privatizados, não observamos Y_{it}^N após a privatização (definido como a receita comercial caso ele não tivesse sido privatizado). Em outras palavras, não sabemos como as receitas teriam evoluído na ausência da privatização. O grande desafio de uma avaliação de impacto é, justamente, encontrar um método robusto de estimar Y_{it}^N .

Como podemos observar na Tabela 5, uma comparação simples das receitas pós-privatização dos aeroportos privados em 2012 e dos demais aeroportos pode refletir não somente o impacto da privatização, mas também outras características que afetam a evolução das receitas comerciais desses aeroportos. Isso porque os aeroportos privatizados são bastante diferentes da média dos demais aeroportos.

Tabela 5 – Características pré-privatização dos 25 maiores aeroportos operados pela Infraero (média 2004-2012)

	Controles	Brasília	Guarulhos	Viracopos
Receita Comercial	15.735.090	40.370.380	224.016.700	22.205.140
Passageiros domésticos	3.097.538	10.988.127	11.445.298	3.072.238
Passageiros internacionais	65.835	190.458	8.517.123	26.581
Aeronaves	38.553	121.491	181.630	46.027
Cargas	22.200.000	74.000.000	391.000.000	211.000.000

Fonte: Elaboração própria com base nos dados da Infraero.

O controle sintético contorna esse problema comparando a evolução das receitas comerciais pós-privatização de cada aeroporto privatizado com a de uma combinação ponderada de aeroportos não-privatizados, de forma a que suas características se assemelhem a do aeroporto privatizado no período pré-privatização.

Para melhor ilustrar como é feita a construção desse aeroporto sintético, suponha que Y_{it}^N possa ser estimado pela seguinte equação:

$$Y_{it}^N = \beta_t X_i + \lambda_t \mu_i + \theta_t + \epsilon_{it} \quad (4)$$

onde X_i é um vetor de variáveis que determinam as receitas comerciais (com seus parâmetros β associados), μ é um vetor de efeitos específicos dos aeroportos que fazem parte do *pool* de doadores (com seus parâmetros λ associados) e ϵ representa o termo de erro, contendo choques idiossincráticos.

Considere $W = (w_2, \dots, w_{j+1})$ um vetor ($J \times 1$), tal que $w_j \geq 0$ para $j=2, \dots, J+1$ e $w_2 + \dots + w_{j+1} = 1$. Note que W é o vetor dos pesos atribuídos para cada aeroporto do *pool* de doadores. Assim, cada valor do vetor W representa uma possível de combinação de pesos para um “aeroporto sintético”, ou seja, uma média ponderada dos aeroportos no *pool* de doadores. Note, ainda, que, conforme (4), o valor das receitas comerciais para cada controle sintético indexado por W será dado por:

$$\sum_{j=2}^{J+1} w_j Y_{jt} = \theta_t + \beta_t \sum_{j=2}^{J+1} w_j X_j + \lambda_t \sum_{j=2}^{J+1} w_j \mu_j + \sum_{j=2}^{J+1} w_j \epsilon_{jt} \quad (5)$$

Suponha que exista um vetor $(w_2^*, \dots, w_{J+1}^*)$ cuja soma seja igual a 1 tal que:

$$\sum_{j=2}^{J+1} w_j^* Y_{j1} = Y_{11}, \dots, \sum_{j=2}^{J+1} w_j^* Y_{jT_0} = Y_{1T_0} \quad e \quad \sum_{j=2}^{J+1} w_j^* X_j = X_1 \quad (6)$$

Os autores demonstram, então, que, atendidas certas condições, é possível encontrar uma combinação ponderada de controles, ou seja, um vetor de pesos W , tal que

$$\hat{\alpha}_{it} = Y_{it}^P - \sum_{j=2}^{J+1} w_j^* Y_{jt} \quad (7)$$

seja um estimador de α_{it} para $t \in \{T_{0+1}, \dots, T\}$.

Nesse caso, poderemos usar $\sum_{j=2}^{J+1} w_j^* Y_{jt}$ como estimador de Y_{it}^N , ou seja, como contrafactual para como as receitas comerciais teriam evoluído caso a privatização não

tivesse ocorrido. Uma prova matemática completa pode ser encontrada em Abadie, Diamond, & Hainmueller (2010).

Note que, por (6), assumimos que existe um vetor W , ou seja, uma combinação de aeroportos do *pool* de doadores tal que a receita comercial ponderada seja igual à receita comercial do aeroporto privatizado em todos os anos pré-privatização e que tenha as mesmas características determinantes das receitas comerciais. Na verdade, não há expectativa de que exista um vetor W tal que se verifiquem essas igualdades, mas que exista um vetor W tal que esses valores sejam muito próximos.

Nesse sentido, vamos escolher o vetor W de forma a minimizar o Erro Quadrado Médio do Estimador – *Mean Squared Prediction Error* (MSPE) da diferença entre a receita comercial do aeroporto privatizado e a do aeroporto sintético nos anos pré-privatização, ou seja, vamos escolher a combinação ponderada de aeroportos do *pool* de doadores de forma a minimizar

$$\|Y_1 - Y_0W\|_V = \sqrt{(Y_1 - Y_0W)'V(Y_1 - Y_0W)} \quad (8)$$

onde V é uma matriz ($k \times k$) simétrica e positiva semidefinida.

Como destacam Abadie, Diamond, & Hainmueller (2010), uma fonte de incerteza em estudos de caso comparativos diz respeito à capacidade do grupo controle de reproduzir o contrafactual, ou seja, como a unidade tratada teria evoluído na ausência do tratamento. No caso específico, a incerteza está relacionada ao fato de não sabermos se os aeroportos escolhidos na ponderação para criação do “aeroporto privatizado sintético” são realmente capazes de reproduzir o que teria ocorrido com o aeroporto privatizado caso a privatização não tivesse ocorrido.

Para superar esse problema, utilizaremos uma série de técnicas de inferência estatística, similares a testes de permutação, propostos em Abadie & Gardeazabal, (2003) e Abadie, Diamond, & Hainmueller (2010). As técnicas envolvem a realização de uma série de testes de placebo “no tempo” (aplicar o tratamento em datas diferentes a que ele, de fato, ocorreu) e “no espaço” (aplicar o tratamento a unidades não tratadas). Essas técnicas serão analisadas em detalhes na seção 1.4.

1.3.2. Bases de Dados

Para realização de nossa análise, construímos um painel de dados balanceado que compreende o período de 2004 a 2016 para os 25 aeroportos da Infraero com maior movimento de passageiros em 2011 (ano da primeira rodada de privatizações).

Em seguida, realizamos as estimações considerando a privatização dos 6 aeroportos entre 2011 e 2014. Para os aeroportos de Brasília, Guarulhos e Viracopos, consideramos 2013 como o primeiro ano de tratamento (ou seja, o primeiro ano em que os aeroportos estiveram efetivamente sob administração privada). Com isso, temos 4 períodos pós-intervenção para observar o impacto da privatização nesses aeroportos. Já para os aeroportos de Confins, Galeão e Natal, consideramos 2015 como o primeiro ano de tratamento. Logo, nesse caso, temos apenas 2 períodos pós-intervenção.

Como nossa lista totaliza 25 aeroportos, sendo que seis deles foram privatizados, cada estimativa conta com 1 aeroporto tratado e 19 aeroportos no *pool* de doadores (os aeroportos privatizados foram excluídos do *pool*). Dessa forma, todas as vezes que nos referimos ao *pool* de doadores estamos nos referindo ao conjunto dos 19 maiores aeroportos não privatizados do País e operados pela Infraero. O *pool* é o mesmo para os 6 aeroportos privatizados. A diferença é somente no período pré-privatização, que compreende os anos de 2004 a 2012 para Brasília, Guarulhos e Viracopos e de 2004 a 2014 para Confins, Galeão e Natal.

Nossa variável de interesse são as receitas comerciais. Assim, os aeroportos de comparação que formam cada aeroporto sintético são selecionados por um algoritmo com base na sua similaridade ao aeroporto tratado antes do tratamento, tanto com relação as variáveis de controle quanto com relação à evolução pretérita das receitas comerciais.

Os dados financeiros relativos aos aeroportos da Infraero foram disponibilizados publicamente como parte do pacote de informações distribuídas a interessados no processo de concessão mais recente, realizado em março de 2017, e que envolveu os aeroportos de Fortaleza, Salvador, Florianópolis e Porto Alegre. Já os dados financeiros referentes aos aeroportos concedidos foram obtidos diretamente do site das Concessionárias e em consulta à Agência Nacional de Aviação Civil – Anac.

Os dados relativos às variáveis de controle foram coletados diretamente do site da Infraero e dos aeroportos concedidos. Em todas as especificações, adotamos as seguintes variáveis de controle: passageiros domésticos, passageiros internacionais, aeronaves e cargas. Trata-se dos principais *drivers* de receitas em aeroportos.

É normal na literatura de aplicação do controle sintético a inclusão de versões defasadas ou da média da variável de interesse no período pré-tratamento, como forma de se diminuir o MSPE (ou seja, de aumentar a qualidade do pareamento). Vários trabalhos sugerem, até mesmo, a inclusão de todas as defasagens anuais da variável de interesse pré-tratamento no modelo estimado²⁴. Kaul et al. (2015), contudo, alertam contra esse procedimento, mostrando que a inclusão de todas as defasagens pode, no limite, tornar as demais variáveis irrelevantes. Os autores recomendam, assim, restringir o uso das defasagens e conduzir estimações alternativas.

De forma a evitar os contratempos apontados por Kaul et al. (2015) e seguindo a recomendação de Ferman, Pinto, & Possebom (2017), optamos por estimar seis especificações distintas. Na primeira, incluímos somente uma defasagem da variável de interesse, relativa ao último ano anterior ao tratamento. Realizamos, ainda, cinco estimações adicionais, incluindo entre os controles: i) a evolução média da variável de interesse no período pré-privatização (modelo 2); ii) somente as variáveis de controle, sem qualquer defasagem da variável de interesse (modelo 3); iii) todas as defasagens da variável de interesse (modelo 4); iv) as defasagens em anos pares (modelo 5); e v) as defasagens em anos ímpares (modelo 6). Os resultados dessas estimações adicionais também foram incluídos nas principais figuras da seção 1.4 e confirmam os resultados do modelo principal²⁵.

1.4 Análise dos Resultados

1.4.1. Aeroportos Sintéticos: resultados das estimações

²⁴ Cavallo et al. (2013), por exemplo, afirmam que a inclusão de todas as defasagens da variável de interesse parece uma escolha óbvia.

²⁵ Para Brasília, Viracopos, Natal e Confins, pareamos cada aeroporto com sua contraparte sintética utilizando como variável de interesse as receitas comerciais. No caso do Aeroporto de Guarulhos e do Galeão, por se tratarem dos dois maiores de nossa amostra, foi necessário modificar ligeiramente a variável de interesse, definindo-a como a evolução das receitas comerciais. Para isso, normalizamos a variável de interesse, estabelecendo a receita comercial como 1 no ano da privatização (2012, no caso de Guarulhos, e 2014, no caso do Galeão). Nesse caso, o modelo de estimação principal não inclui qualquer defasagem da variável de interesse (modelo 3). Para facilitar a comparação do impacto da privatização nos seis aeroportos, os gráficos são todos apresentados considerando a evolução das receitas comerciais.

Como explicado anteriormente, para cada aeroporto privatizado, construímos um aeroporto sintético, a partir da combinação convexa de aeroportos não-privatizados. O peso de cada aeroporto do *pool* de doadores na criação de cada aeroporto sintético pode ser observado na Tabela 6. A tabela indica, por exemplo, que a evolução da receita comercial em Viracopos entre 2004 e 2012 é melhor estimada por uma combinação dos aeroportos de Santos Dumont (RJ), Porto Alegre e Foz do Iguaçu (os demais aeroportos recebem peso zero). Já no caso de Brasília, os aeroportos que compõem o aeroporto sintético foram Congonhas (SP) e Porto Alegre.

Tabela 6 – Peso dos Aeroportos do *Pool* nos Aeroportos Sintéticos

Brasília		Guarulhos		Viracopos	
Congonhas (SP)	42%	Congonhas (SP)	64%	Santos Dumont (RJ)	54%
Porto Alegre	58%	Manaus	36%	Porto Alegre	16%
				Foz do Iguaçu	30%
Natal		Galeão		Confins	
Porto Alegre	2%	Congonhas (SP)	49%	Santos Dumont (RJ)	36%
Salvador	2%	Salvador	24%	Porto Alegre	53%
Fortaleza	14%	Manaus	27%	Salvador	5%
Foz do Iguaçu	82%			Foz do Iguaçu	6%

Fonte: Elaboração própria com base nos resultados estimados.

A Tabela 7 traz a comparação entre as características observáveis pré-privatização dos aeroportos privatizados, do *pool* de doadores e dos aeroportos sintéticos. Nota-se que as médias do *pool* de doadores diferem substancialmente dos aeroportos privatizados na maior parte das características. Por outro lado, os aeroportos sintéticos apresentam grau significativamente maior de semelhança com os aeroportos reais. As exceções são o tráfego internacional de passageiros em Guarulhos e Galeão e a movimentação de cargas em Guarulhos e Viracopos. Isso decorre do fato de Guarulhos e Galeão representarem mais de 95% do tráfego internacional de passageiros no País no período, inexistindo, assim, no *pool* de doadores aeroportos que pudessem mimetizar essa característica. O mesmo ocorre no caso da movimentação de cargas em Guarulhos e Viracopos. Ainda nesses casos, contudo, os aeroportos sintéticos oferecem um pareamento superior à média dos aeroportos no *pool* de doadores.

Tabela 7– Características pré-privatização - média 2004-2012 para o primeiro grupo e 2004-2014 para o segundo grupo (em milhares)

	Controles	Brasília		Guarulhos		Viracopos	
		Real	Sintética	Real	Sintético	Real	Sintético
Passageiros Domésticos	3.097	10.988	8.840	11.445	9.913	3.072	3.981
Passageiros Internacionais	66	191	187	8.517	482	27	58
Aeronaves	39	122	101	182	113	46	48
Cargas	22.200	74.000	39.100	391.000	78.700	211.000	84.400

	Controles	Natal		Galeão		Confins	
		Real	Sintético	Real	Sintético	Real	Sintético
Passageiros Domésticos	2.894	1.527	1.651	9.010	9.384	5.505	5.379
Passageiros Internacionais	59	64	59	2.853	106	204	204
Aeronaves	35	18	17	115	104	70	63
Cargas	20.000	7.496	7.452	96.900	70.000	20.300	20.500

Fonte: Elaboração própria com base nos dados da Infraero e nas estimações realizadas.

Apresentamos os resultados da aplicação do método do controle sintético na Figura 1 e na Figura 2 a seguir. A primeira reproduz a evolução dos aeroportos reais e sintéticos considerando o modelo principal. A segunda traz a diferença entre a evolução real e a evolução sintética nas seis especificações estimadas.

Como se nota, para a maioria dos aeroportos privatizados, os aeroportos sintéticos foram capazes de reproduzir com precisão a evolução das receitas comerciais no período pré-privatização. A única exceção foi o Aeroporto de Confins, cujo pareamento não foi realizado com êxito. Não surpreende que assim seja, já que se trata de um aeroporto com características bastante peculiares²⁶. É natural, portanto, que nenhuma combinação de

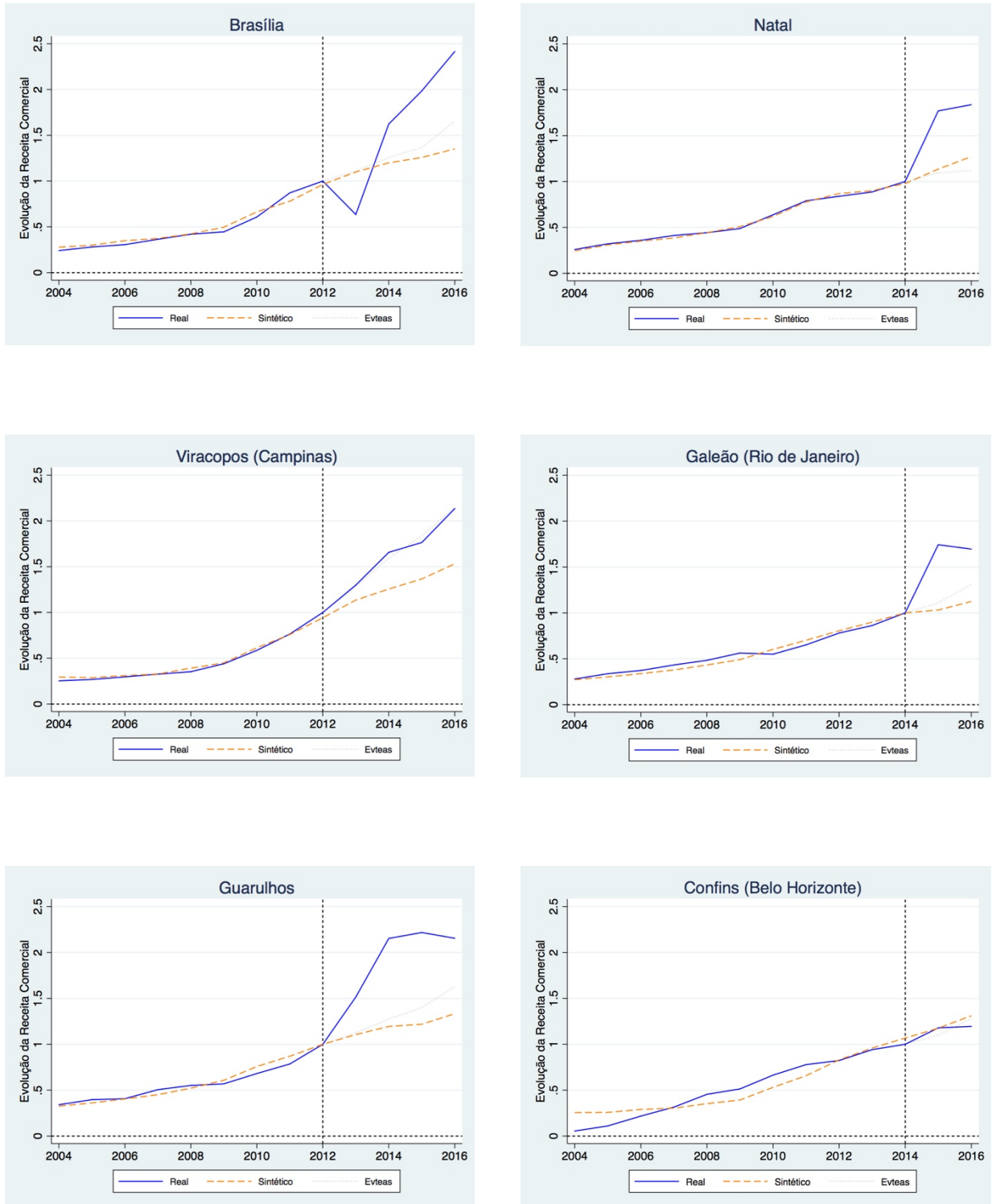
²⁶ Pelo menos até 2009, Confins apresentava um movimento muito pequeno, sendo a região de Belo Horizonte atendida, majoritariamente, pelo Aeroporto de Pampulha (localizado no centro da cidade). A partir de então, o estado começa um programa de estímulo ao Aeroporto de Confins, limitando significativamente o acesso das empresas ao Aeroporto de Pampulha, e redirecionando voos para Confins, o que gera um crescimento significativo das operações daquele aeroporto.

aeroportos do *pool* de doadores consiga reproduzir razoavelmente as características de Confins no período pré-privatização. Por essa razão, seguindo a recomendação de Abadie, Diamond, & Hainmueller (2010), não consideraremos os resultados do estimador de controle sintético para este aeroporto (os autores recomendam que o estimador não seja usado se o ajuste pré-tratamento for fraco).

No caso dos demais aeroportos, para os quais a evolução das receitas comerciais pré-tratamento é mimetizada com precisão pelos aeroportos sintéticos, o impacto estimado da privatização nas receitas comerciais é dado simplesmente pela diferença entre as duas linhas no período pós-privatização – ou seja, pela diferença entre a evolução real das receitas comerciais e a evolução sintética no período pós-privatização. A análise a seguir leva em consideração somente os aeroportos nos quais o pareamento foi realizado com êxito.

O impacto da privatização mostrou-se de grande magnitude para os cinco aeroportos. A título de exemplo, observamos que, dois anos após a privatização, as receitas comerciais de Viracopos eram 66% maiores do que eram na época da privatização, enquanto nossas estimações indicam que elas seriam apenas 26% maior no cenário contrafactual (ou seja, caso esse aeroporto continuasse administrado pela Infraero). No caso de Natal, já no ano seguinte ao início da operação pela Concessionária privada, as receitas comerciais subiram 77%, o que representa um aumento significativo quando comparado ao crescimento contrafactual estimado de apenas 14%.

Figura 1 - Evolução real e contrafactual das receitas comerciais: Aeroportos Reais vs. Aeroportos Sintéticos



Os desvios entre a evolução real e sintética das receitas comerciais podem ser melhor observados na Figura 2, que traz a diferença entre a evolução real das receitas comerciais e a evolução contrafactual - ou seja, o impacto da privatização. Em primeiro lugar, nota-se que, na grande maioria das especificações, no período pré-privatização, os aeroportos sintéticos reproduziram bastante bem o comportamento dos aeroportos reais, com desvios muito pequenos.

Já no período pós-privatização, observa-se o contrário: a trajetória dos aeroportos descola-se significativamente da trajetória dos sintéticos. Chama a atenção a consistência dos resultados estimados, sendo os efeitos estimados bastante semelhantes em todas as especificações dos cinco aeroportos. Destaca-se, ademais, o fato de os efeitos estimados serem relativamente próximos, mesmo para aeroportos distintos como Galeão e Natal.

O gráfico de Brasília apresenta um movimento peculiar no primeiro ano pós-privatização, com redução significativa das receitas. Isso é explicado pelo fato de que, nesse aeroporto, os investimentos obrigatórios implicaram a reforma de boa parte das instalações onde se localizavam as lojas, restaurantes e outros estabelecimentos comerciais do aeroporto e, conseqüentemente, o fechamento temporário dessas instalações. Essa tendência de redução, contudo, é rapidamente invertida, já a partir do segundo ano da concessão.

Os impactos em termos absolutos da privatização nos cinco aeroportos podem ser observados na Tabela 8. A análise dos dados da Tabela deixa evidente os resultados extremamente significativos da privatização nas receitas dos aeroportos. Como vemos, já no segundo ano pós-privatização, há um ganho de mais de R\$ 400 milhões no aeroporto de Guarulhos; quando consideramos os ganhos acumulados nos quatro anos posteriores, o impacto da privatização chega a mais de R\$ 1,3 bilhão. Impacto semelhante em termos relativos se verifica no Aeroporto de Brasília, onde as receitas comerciais aumentaram mais de 140% em apenas 4 anos. Pelas nossas estimativas, esse aumento seria de 40% caso o aeroporto continuasse sob operação da Infraero, o que se traduz em um impacto anual de cerca de R\$ 100 milhões nas receitas comerciais no quarto ano da privatização.

A privatização gerou, portanto, um ganho operacional imediato e extremamente significativo nos cinco aeroportos, o que demonstra a existência de grande diferencial de eficiência na operação de aeroportos entre a Infraero e as concessionárias privadas.

Figura 2 – Impacto estimado da privatização nas receitas comerciais nas 6 especificações

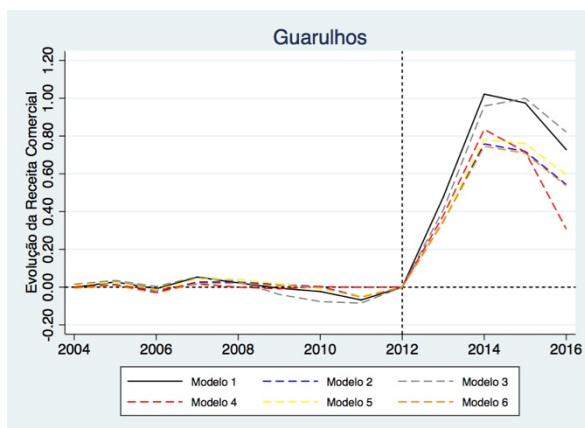
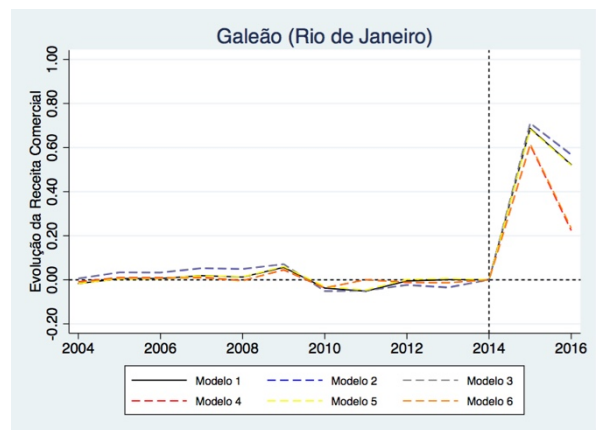
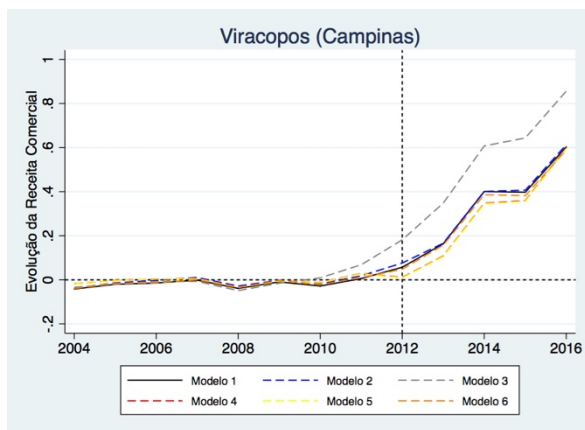
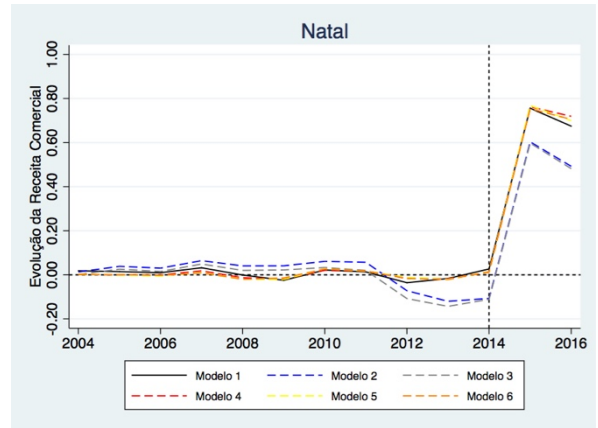
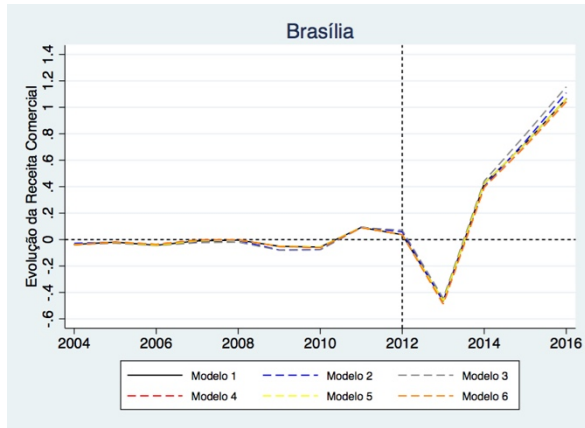


Tabela 8 – Impacto Anual Estimado da Privatização na Evolução das Receitas Comerciais (em milhões de R\$)

	Brasília			Guarulhos			Viracopos		
	Real	Sintético	Impacto	Real	Sintético	Impacto	Real	Sintético	Impacto
Ano 0	91	88	3	423	423	0	47	44	3
1º Ano	58	100	-42	641	468	173	61	53	8
2º Ano	148	110	38	911	505	405	77	59	18
3º Ano	181	115	66	938	515	422	82	64	18
4º Ano	221	124	97	911	564	347	100	71	28
Impacto Total Pós-Privatização			159			1347			72

	Natal			Galeão		
	Real	Sintético	Impacto	Real	Sintético	Impacto
Ano 0	15	15	0	245	245	0
1º Ano	27	17	10	427	253	174
2º Ano	28	19	9	415	276	139
Impacto Total Pós-Privatização			19			313

Fonte: Elaboração própria com base nos resultados estimados.

Na Figura 1, incluímos, ainda, para efeitos de comparação, a evolução esperada nas receitas comerciais tendo por base os Estudos de Viabilidade Técnica, Econômica e Ambiental - EVTEAs do governo (linha pontilhada). Como vimos na seção 1.3, trata-se dos estudos em que o governo realiza o *valuation* dos aeroportos e que são utilizados como base para a definição do valor mínimo do leilão. Curiosamente, na maioria dos casos, os EVTEAs acompanharam a evolução do aeroporto sintético, ou seja, a evolução das receitas comerciais caso esses aeroportos continuassem a ser operados pela Infraero. Nesse sentido, os estudos mostraram-se bastante conservadores. A única exceção foi o aeroporto de Viracopos. Nesse caso, os EVTEAs reproduziram com exatidão a evolução real das receitas comerciais, o que pode explicar o fato de o ágio deste aeroporto ter sido o menor entre os cinco aeroportos avaliados.

1.4.2. Inferência: Testes de Placebo e Falsificações

Nesta seção, conduziremos uma série de testes de placebo, com vistas a assegurar a significância estatística e a robustez dos resultados estimados.

Inicialmente, realizaremos os chamados “testes de placebo no espaço”. Para isso, aplicaremos o método de controle sintético a todos os aeroportos do *pool* de doadores, como se cada um deles tivesse sido objeto de uma privatização. A ideia é calcular a probabilidade de que impactos tão extremos como os estimados para os aeroportos privatizados possam ser obtidos por acaso. Em outras palavras, queremos avaliar se um ou mais aeroportos não-privados apresentaram uma evolução semelhante no período pós-privatização, ainda que não tenham sido privatizados.

A Figura 3 compara os impactos da privatização na evolução das receitas comerciais nos aeroportos privatizados e dos tratamentos placebos no *pool* de doadores. Como se nota, tanto no caso de Brasília, Guarulhos e Viracopos como no caso de Galeão e Natal, a evolução dos aeroportos privatizados é bastante superior aos demais. Isso ocorre mesmo considerando os aeroportos cujo ajuste pré-tratamento não é tão bom.

Seguindo o método desenvolvido por Abadie, Diamond, & Hainmueller (2010), podemos ranquear os resultados dos aeroportos e calcular o p-valor implícito como a probabilidade de obter uma estimativa tão grande quanto a obtida quando o tratamento é designado aleatoriamente no conjunto de todas as observações. O p-valor implícito é dado, portanto, simplesmente pelo número de aeroportos que apresentam um “impacto” superior ao dos aeroportos privatizados no período pós-privatização. Note que, como nossas amostras são compostas por 20 aeroportos (1 privatizado e 19 controles), o menor valor assumido por p é 0,05, o que ocorre no caso de não haver nenhum aeroporto no *pool* de doadores que apresenta um “impacto” superior do tratamento placebo, reduzindo a fórmula a $(1/20 = 0,05)$. Esse é justamente o caso dos cinco aeroportos privatizados, como se pode conferir, visualmente, na Figura 3.

Outra forma de avaliar a qualidade do pareamento é comparar a razão do MSPE pós e pré intervenção para todos os aeroportos da nossa amostra. Trata-se de método mais robusto, uma vez que considera a qualidade do ajuste pré-intervenção para todos os aeroportos. A intuição por trás dessa abordagem é averiguar para quais aeroportos o nível de distanciamento do aeroporto real com relação ao aeroporto sintético no momento pós-tratamento é maior quando comparado ao nível de distanciamento pré-intervenção. Novamente, queremos avaliar a probabilidade de impactos de magnitude semelhante aos observados nos aeroportos privatizados possam ser resultados do acaso. Os resultados podem ser visualizados na Figura 4.

Figura 3 – Diferença da evolução das receitas comerciais em Aeroportos Privatizados e Placebos (20 aeroportos)

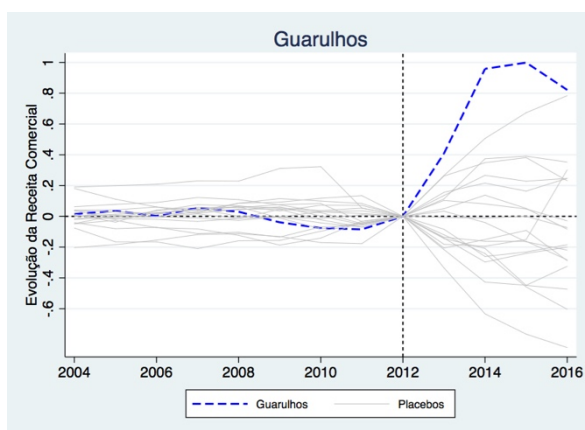
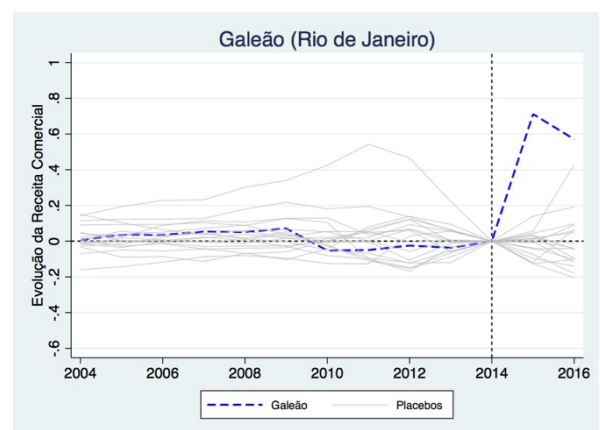
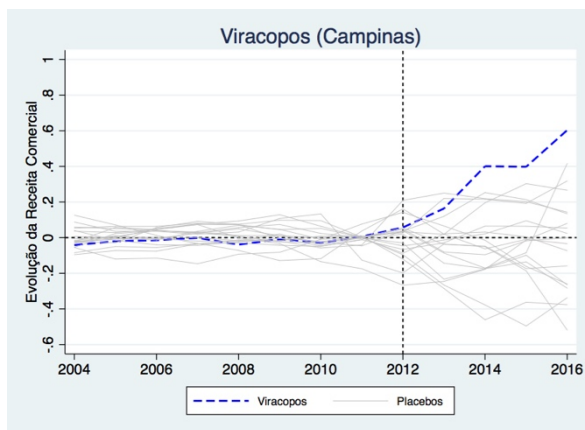
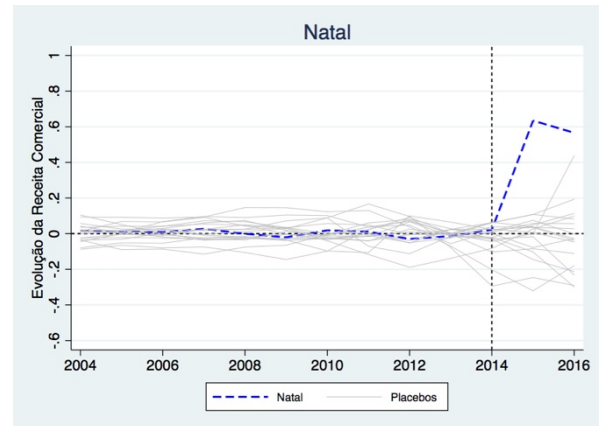
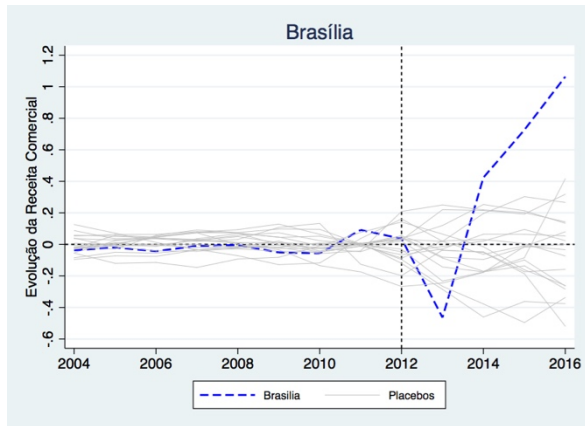
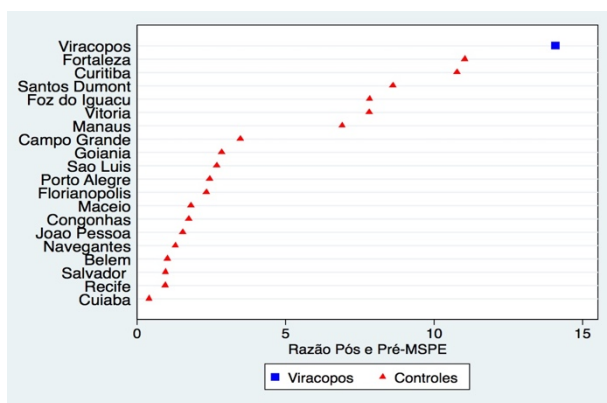
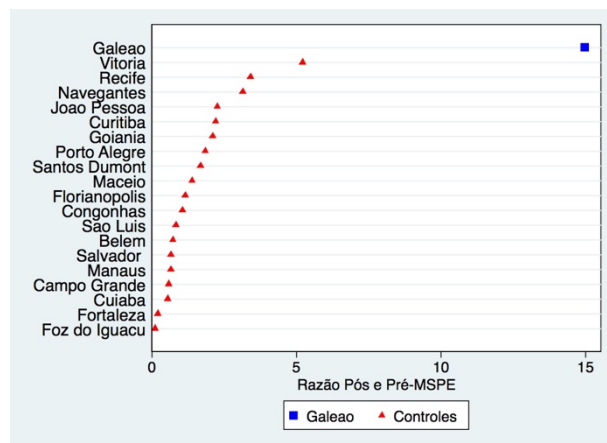
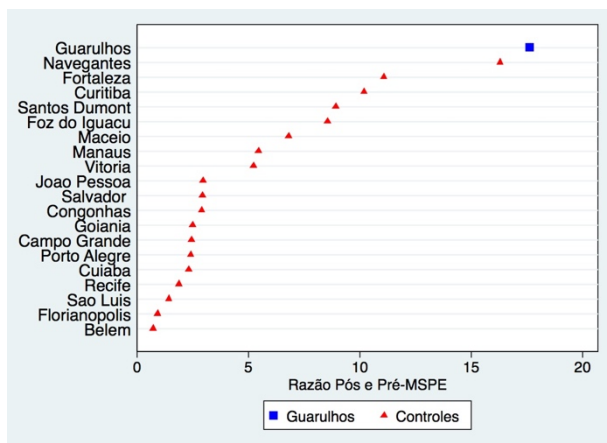
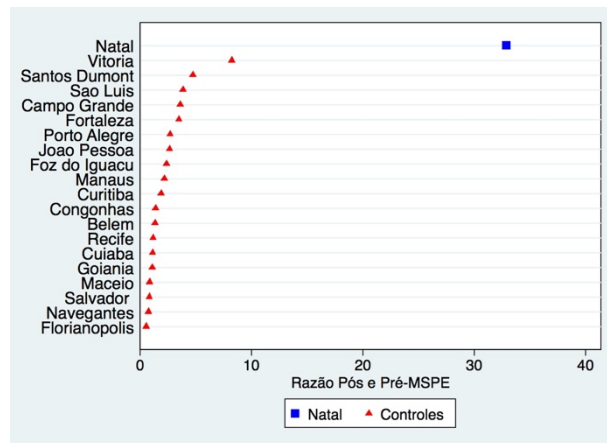
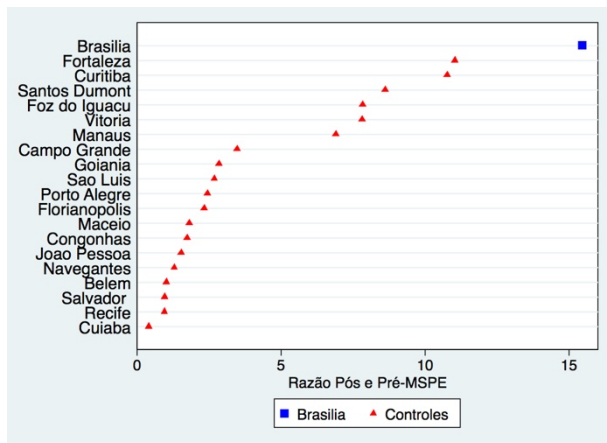


Figura 4 – Razão do MSPE Pós-privatização e Pré-Privatização (Aeroportos privatizados e controles)



Mais uma vez, os valores dos aeroportos privatizados aparecem em destaque. A razão do MSPE pós e pré-privatização chega, por exemplo, a 16 em Brasília, enquanto o maior valor observado para um aeroporto do *pool* de doadores é de 11 e a média é de cerca de 4. Já para Natal, por exemplo, os resultados são ainda mais expressivos: enquanto a razão observada para esse aeroporto é de 32, o maior valor observado no grupo de controle é de 8 e a média é de 2. Para os cinco aeroportos analisados, a razão do MSPE pós e pré-privatização foi a maior observada. Utilizando raciocínio semelhante ao dos parágrafos anteriores, novamente temos um p-valor implícito de 0,05 para os cinco aeroportos. Isso demonstra que os aeroportos privatizados foram aqueles que tiveram o maior descolamento da evolução real das receitas comerciais em relação à evolução sintética no período pós-privatização quando comparado com o período pré-privatização, diminuindo, assim, a probabilidade de que os resultados estimados tenham ocorrido por acaso.

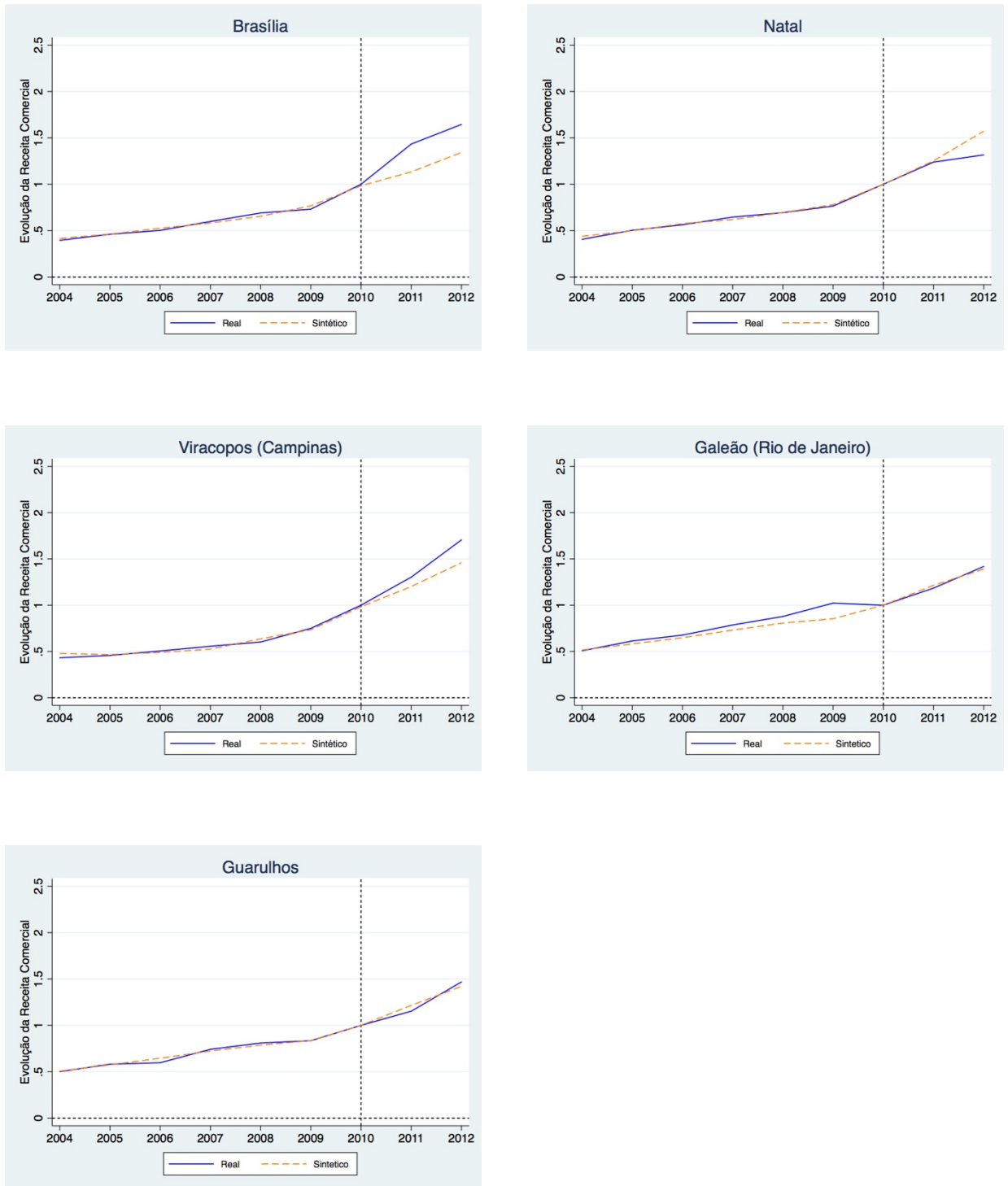
Com isso, concluímos os testes de placebo “no espaço”, que demonstram a robustez do resultado estimado.

Analisaremos agora os testes “no tempo”, o que implica estimar o impacto da privatização em períodos distintos ao que ela, de fato, ocorreu. A Figura 5 traz o resultado das estimações de um tratamento placebo para os cinco aeroportos em 2010.

Como se nota, também no caso dos placebos “no tempo”, os resultados mostram-se robustos, não havendo qualquer alteração relevante na evolução das receitas de Guarulhos, Galeão, Natal e Viracopos no período pós-2012. No caso de Brasília, vemos uma pequena alteração no período pós-2010, que, contudo, é bastante inferior aos resultados estimados para a privatização.

Finalmente, como argumentam Ferman, Pinto, & Possebom (2017), a ausência de uma regra clara na literatura sobre como selecionar as variáveis utilizadas para o pareamento do controle sintético, principalmente no que tange à inclusão ou não de defasagens da variável dependente, deixa o modelo suscetível à busca pelo pesquisador de uma especificação que apresente os resultados desejados (*cherry picking*). Para endereçar essa preocupação, é importante demonstrar que o resultado é robusto a diferentes especificações do modelo.

Figura 5 - Evolução real e contrafactual das receitas comerciais: tratamento placebo em 2010



Como explicamos ao final da seção 1.3, realizamos cinco estimações adicionais, além daquela cujos resultados foram expostos no corpo principal deste trabalho. Apresentamos, na Tabela 9, as estimações de p-valor e do R² das seis estimações para cada um dos aeroportos²⁷. Os indicadores da tabela mostram bastante coerência entre os impactos estimados nas diversas especificações e confirmam os resultados principais expostos nesta seção.

Tabela 9 – Comparação entre Modelos

	Brasília		Guarulhos		Viracopos	
	p-valor	R ²	p-valor	R ²	p-valor	R ²
Modelo 1	0.05	96.7%	0.05	97.4%	0.05	98.5%
Modelo 2	0.05	95.3%	0.05	98.6%	0.05	98.3%
Modelo 3	0.05	94.7%	0.05	94.4%	0.10	92.1%
Modelo 4	0.10	96.7%	0.05	99.8%	0.05	99.5%
Modelo 5	0.05	96.7%	0.15	98%	0.05	99.5%
Modelo 6	0.10	96.7%	0.10	98.6%	0.10	98.8%

	Natal		Galeão	
	p-valor	R ²	p-valor	R ²
Modelo 1	0.05	99.4%	0.05	98.4%
Modelo 2	0.05	94.9%	0.05	94.9%
Modelo 3	0.05	94.6%	0.05	94.6%
Modelo 4	0.05	99.7%	0.05	98.9%
Modelo 5	0.05	99.7%	0.05	98%
Modelo 6	0.05	99.7%	0.05	98.9%

Fonte: Elaboração própria nas estimações conduzidas.

²⁷ O R² foi calculado conforme proposto por Ferman, Pinto, & Possebom (2017), ou seja:

$$\bar{R}^2 = 1 - \frac{\sum_{t=1}^{T_0} (Y_{i,t} - \hat{Y}_{1,t}^N)^2}{\sum_{t=1}^{T_0} (Y_{i,t} - \bar{Y}_1)^2}$$

1.5 Discussão dos Resultados e Conclusão

Este trabalho teve como objetivo contribuir para a literatura de avaliação de políticas públicas no País, realizando a primeira avaliação de impacto rigorosa da política brasileira de privatização de aeroportos.

Como vimos, os impactos da privatização na geração de receitas comerciais mostraram-se estatisticamente e economicamente bastante significativos em todos os aeroportos avaliados. Para os três aeroportos privatizados em 2012 (Brasília, Guarulhos e Viracopos), em apenas quatro anos, as receitas comerciais aumentaram entre 113% e 132%, enquanto, pelas nossas estimativas, teriam crescido entre 33% e 51% caso esses aeroportos continuassem sob operação da Infraero. Isso representa um ganho de receitas comerciais de quase R\$ 500 milhões somente em 2016. Já para os dois aeroportos cuja privatização teve início em 2015, o aumento das receitas nos dois primeiros anos pós-privatização alcançou 70%-87% enquanto teria sido de apenas 13%-26% com a Infraero – o que representou um ganho de receitas de cerca de R\$ 150 milhões em 2016.

Conseqüentemente, os resultados estimados nesse trabalho levam-nos a concluir pela maior probabilidade de ocorrência do que chamamos de “hipótese positiva”, ou seja, de que os ágios no leilão decorram de um grande diferencial de eficiência na operação dos aeroportos por parte de empresas privadas. Em termos simples, as empresas privadas têm se mostrado muito mais eficientes do que a Infraero na operação de aeroportos.

Os resultados estimados constituem, ainda, forte argumento em favor da continuidade da política de privatização do setor aeroportuário nacional. Nunca é demais lembrar que os valores atualmente arrecadados a título de outorga por cada aeroporto é bastante superior aos lucros que a Infraero vinha obtendo em sua operação (vide Tabela 4). Além disso, o modelo de regulação por *price-cap* garante que as tarifas aeroportuárias não serão reajustadas a valores acima da inflação, assegurando a modicidade tarifária para os passageiros (vide seção 1.2.2).

Nesse sentido, a política de privatização parece beneficiar os mais diversos segmentos sociais: ganha o poder público, que arrecada muito mais recursos desses ativos (tanto com outorga, como com impostos); ganha o setor privado, que passa a auferir lucros com um setor antes completamente estatizado; ganham os passageiros, que obtém um servido de qualidade muito superior - como mostra a maior disposição a pagar dos agentes por serviços não aeronáuticos - ao mesmo tempo em que arcam com um custo igual ou inferior pelos serviços essenciais –, uma vez que as tarifas aeronáuticas foram reguladas ao preço vigentes no momento da concessão; e ganha o restante sociedade que,

em tese, irá usufruir, por meio de políticas públicas, do aumento de arrecadação governamental. Nesse processo, os únicos prejudicados parecem ser os grupos de interesse que transitam ao redor da Infraero (ex. partidos políticos com cargo na empresa, fornecedores de longa data, etc.).

A existência de um grande diferencial entre as empresas privadas e a Infraero é reforçada, ainda, pela comparação entre a evolução observada nas receitas comerciais e a evolução prevista pelos EVTEAs. Como vimos, em praticamente todos os casos, os EVTEAs mimetizaram o que teria ocorrido caso o aeroporto continuasse sob operação da Infraero²⁸. Nesse sentido, é natural que os ágios verificados tenham sido significativos, uma vez que a evolução real das receitas é bem superior à prevista nos estudos de viabilidade.

É importante ressaltar que isso não reflete, necessariamente, uma falha dos EVTEAs. Estudos de viabilidade que amparam processos de concessão devem ser, de fato, conservadores. Não se deve perder de vista que o objetivo principal desses estudos é chegar a um “valor mínimo razoável” exigido pela privatização desses ativos, além de realizar uma avaliação do atual estado da infraestrutura. O governo não deve pretender acertar exatamente o valor do ativo (até porque, por se tratar de um exercício preditivo a longo prazo, ele naturalmente terá grandes doses de incerteza). Afinal, caso os estudos subestimem o potencial de geração de receitas (e de economia de custos) de empresas privadas na operação de ativos privatizados, como parece ter ocorrido no caso dos aeroportos, o mecanismo do leilão atua, justamente, no sentido de corrigir essa subestimação. Por outro lado, caso os estudos superestimem o potencial de geração de receitas (e de economia de custos), o leilão tenderá a ser vazio, o que compromete a privatização dos ativos.

Ao mesmo tempo, é fundamental que o agente público tenha consciência de que essas estimativas são conservadoras. Particularmente no atual momento do setor aeroportuário, esse conhecimento é de extrema relevância. Isso porque, até o presente momento, o governo federal concedeu à iniciativa privada apenas aeroportos de médio e grande porte, cuja viabilidade econômica era praticamente certa. À medida que o processo de privatização avance para aeroportos menores, os EVTEAs elaborados podem

²⁸ Não podemos descartar nessa análise, conforme já alertamos no início desse trabalho, a possibilidade de ter havido um “feito aprendizado” na Infraero, tendo sido também ela indiretamente afetada pelo tratamento, o que levou a um crescimento das receitas comerciais dos demais aeroportos.

passar a impressão de que esses aeroportos não são lucrativos, desestimulando a continuação do processo de privatização. Essa impressão, contudo, pode ser falsa, uma vez que é baseada em estudos de viabilidade econômica que, como vimos, são bastante conservadores.

Uma observação final importante diz respeito à literatura de privatização e regulação em geral. Em Resende, Caldeira, & Fonseca (2016), os autores esboçam uma crítica teórica a trabalhos que buscaram comparar os resultados em termos de bem-estar social de um aeroporto público, privado regulado e privado desregulado²⁹. Como argumentam os autores, trata-se de trabalhos que sofrem da chamada “falácia do Nirvana” de Demsetz, (1969), ou seja, partem de modelos idealmente construídos para buscar distorções no mundo real e concluir que a intervenção estatal – seja na forma de produção direta, seja na forma de regulação – é benéfica, sem analisar adequadamente os custos e distorções inerentes à essa intervenção.

Littlechild (2010) já havia demonstrado a limitação desse tipo de análise. Segundo o autor, a análise de setores industriais nacionalizados no Reino Unido na década de 1970 e 1980 sofria de problema semelhante, na medida em que vários autores buscavam normas de precificação e investimentos eficientes para empresas públicas, enquanto deixavam de lado os problemas mais fundamentais, que tinham natureza dinâmica: não fazia sentido tomar curvas de custo e demanda como dadas para se chegar a uma suposta precificação eficiente; a privatização era recomendada justamente com vistas a alterar a forma dessas curvas. Vimos, neste trabalho, como, particularmente no caso do setor aeroportuário brasileiro, a privatização alterou de forma significativa a curva de receitas dos aeroportos.

Esperamos, assim, haver contribuído para demonstrar a limitação de modelos estáticos na discussão sobre operação pública, regulação e desregulação. Sob o ponto de vista metodológico, a centralidade do conceito de concorrência perfeita na análise de falhas de mercado combinada com o foco na análise estática de equilíbrio pode, em muitos casos, levar a modelagens, no mínimo, limitadas. Isso porque, grande parte desses modelos parte da premissa (muitas vezes implícita) de que a privatização não afetará a curva de custos e de receitas, resumindo a discussão de propriedade pública *versus* privada a um problema de precificação eficiente. Trata-se, contudo, a nosso ver, de uma abordagem metodologicamente equivocada. A discussão sobre operação pública *versus*

²⁹ Exemplos: Basso, 2008; Martín & Pilar Socorro, 2009; e Oum, Zhang, & Zhang, 2004.

privada é uma discussão essencialmente dinâmica e, portanto, não pode ser adequadamente analisada em modelos que desconsiderem essa característica.

Por fim, julgamos importante que o estudo aqui realizado seja complementado por uma avaliação da evolução dos custos de aeroportos privatizados. Como vimos, o diferencial de eficiência de uma empresa privada em relação à Infraero pode aparecer, basicamente, de duas formas: na evolução das receitas comerciais e na evolução dos custos. Dessa forma, uma análise desse diferencial somente será completa com uma análise da evolução dos custos dos aeroportos privatizados.

1.6 Referências Bibliográficas

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493–505. <https://doi.org/10.1198/jasa.2009.ap08746>

Abadie, A., Diamond, A. J., & Hainmueller, J. (2014). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*. . <https://doi.org/10.2139/ssrn.1950298>

Abadie, A., & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, 93(1), 113–132. <https://doi.org/10.1257/000282803321455188>

ACI, A. C. I. (2008). *Airport Economics Survey*.

Athey, S., & Imbens, G. (2016). The State of Applied Econometrics - Causality and Policy Evaluation. Retrieved from <http://arxiv.org/abs/1607.00699>

ATRS, A. T. R. S. (2011). *Global Airport Benchmarking Report*.

Basso, L. J. (2008). Airport deregulation: Effects on pricing and capacity. *International Journal of Industrial Organization*, 26(4), 1015–1031. <https://doi.org/10.1016/j.ijindorg.2007.09.002>

CAA, C. A. A. (2000). *The “Single Till” and the “Dual Till” Approach to the Price Regulation of Airports*. Consultation Paper, London, December.

Cavallo, E., Galiani, S., Noy, I., & Pantano, J. (2013). Catastrophic Natural Disasters and Economic Growth. *Review of Economics and Statistics*, 95(5), 1549–1561. https://doi.org/10.1162/REST_a_00413

Demsetz, H. (1969). Information and Efficiency: Another Viewpoint. *The Journal of Law and Economics*, 12(1), 1–22. <https://doi.org/10.1086/466657>

Doganis, R. (1992). *The airport business*. Routledge.

Ferman, B., Pinto, C., & Possebom, V. (2017). Cherry Picking with Synthetic Controls *. Retrieved from <https://dl.dropboxusercontent.com/u/12654869/FPP - Cherry Picking.pdf>

Forsyth, P. J. (2002). Privatisation and regulation of Australian and New Zealand airports. *Journal of Air Transport Management*, 8(1), 19–28. Retrieved from

http://econpapers.repec.org/article/eeejaitra/v_3a8_3ay_3a2002_3ai_3a1_3ap_3a19-28.htm

Forsyth, P. J. (2003). Regulation under stress: developments in Australian airport policy. *Journal of Air Transport Management*, 9(1), 25–35. Retrieved from http://econpapers.repec.org/article/eeejaitra/v_3a9_3ay_3a2003_3ai_3a1_3ap_3a25-35.htm

Graham, A. (2009). How important are commercial revenues to today's airports? *Journal of Air Transport Management*, 15(3), 106–111. <https://doi.org/10.1016/j.jairtraman.2008.11.004>

Jones, I., Viehoff, I., Marks, P., Jones, I., Viehoff, I., & Marks, P. (1993). The economics of airport slots. *Fiscal Studies*, 14(4), 37–57. Retrieved from http://econpapers.repec.org/article/ifsfistud/v_3a14_3ay_3a1993_3ai_3a4_3ap_3a37-57.htm

Kaul, A., Kloessner, S., Pfeifer, G., & Schieler, M. (2015). Synthetic Control Methods : Never Use All Pre-Intervention Outcomes as Economic Predictors, 1–19. Retrieved from http://www.oekonometrie.uni-saarland.de/papers/SCM_Predictors.pdf

Littlechild, S. (2010). The nature of competition and the regulatory process. Retrieved from <http://www.eprg.group.cam.ac.uk/wp-content/uploads/2014/01/thenature-of-competition-and-the-regulatory-process.pdf>

Martín, J. C., & Pilar Socorro, M. (2009). A new era for airport regulators through capacity investments. *Transportation Research Part A: Policy and Practice*, 43(6), 618–625. <https://doi.org/10.1016/j.tra.2009.04.001>

Oum, T. H., Zhang, A., & Zhang, Y. (2004). Alternative Forms of Economic Regulation and their Efficiency Implications for Airports. *Journal of Transport Economics and Policy*, 38(2), 217–246. Retrieved from http://econpapers.repec.org/article/tpejtecpo/v_3a38_3ay_3a2004_3ai_3a2_3ap_3a217-246.htm

Productivity Commission. (2011). *Economic Regulation of Airport Services*. Report No. 54. Commonwealth of Australia.

Resende, C. C., Caldeira, T., & Fonseca, R. (2016). Aeroportos competem? Revisão da Literatura e Opções Regulatórias Brasileiras. *Revista de Defesa Da Concorrência*, 4(2), 5–44. Retrieved from <http://revista.cade.gov.br/index.php/revistadedefesadaconcorrenca/article/view/250>

Starkie, D. (2001). Reforming UK Airport Regulation. *Journal of Transport Economics*

and Policy, 35(1), 119–135. Retrieved from http://econpapers.repec.org/article/tpejtecpo/v_3a35_3ay_3a2001_3ai_3a1_3ap_3a119-135.htm

Starkie, D. (2008). A critique of the single-till. In *Aviation Markets* (pp. 123–130). Ashgate, Alershot.

Starkie, D., & Yarrow, G. (2000). The Single Till Approach to the Price Regulation of Airports. *CAA. London*.

Ensaio II

2 ENSAIO II: Ampliação da jornada escolar e indicadores sociais: uma avaliação de impacto do Programa Mais Educação no trabalho infantil e no desemprego das mães

Resumo

O Programa Mais Educação - PME é a principal política de estímulo à educação integral atualmente vigente no Brasil. Os efeitos do PME no desempenho acadêmico dos estudantes e no fluxo escolar já foram avaliados em diversos trabalhos. A grande maioria não encontrou impacto significativo do programa nessas variáveis. Neste trabalho, exploramos a descontinuidade existente no critério de elegibilidade do programa a partir de 2012 para avaliar seu impacto em dois indicadores sociais: o trabalho infantil e o desemprego das mães. Além disso, avaliamos o impacto do programa em indicadores relacionados à satisfação dos pais e dos estudantes com a escola, ao comportamento dos estudantes e aos índices de estudantes em recuperação. Não encontramos evidência de impacto significativo do programa em nenhum dos indicadores avaliados. De forma a reforçar os resultados estimados, conduzimos, ainda, uma avaliação dos efeitos do programa por diferenças em diferenças. Este modelo tampouco mostrou qualquer impacto significativo do programa nas variáveis analisadas.

Palavras-chave: Educação; Jornada Escolar; Desemprego; Trabalho Infantil; Avaliação de Impacto.

2.1 Introdução

O Programa Mais Educação é a principal estratégia do Governo Federal de estímulo à implementação da educação integral no país. A defesa da educação integral é, em geral, realizada com base no seu suposto impacto no desempenho acadêmico de crianças e jovens. Infelizmente, com base nesse critério, o Programa Mais Educação não vem apresentando bons resultados, como mostra a análise de diversas avaliações de impacto recentes do programa, tais como: Mendes (2011), Pereira (2011), Xerxenevsky (2012), Fundação Itaú (2015) e Oliveira & Terra (2016). Os resultados desses trabalhos são consistentes no sentido de apontar a ausência de impacto da extensão da jornada escolar financiada pelo PME no desempenho dos estudantes em testes padronizados.

Ainda que tais resultados sejam preocupantes e recomendem um redesenho significativo do Programa, é importante destacar que programas de expansão da jornada escolar podem ter efeitos não apenas no desempenho acadêmico. É natural que a escola também desempenhe um papel não acadêmico na vida das crianças e jovens e daqueles a seu redor. Diversas evidências empíricas suportam esse raciocínio. A título de exemplo, Jacob & Lefgren (2003) encontram impactos significativos e positivos da rotina escolar na redução de certos tipos de crimes, tais como crimes relacionados à propriedade. Kruger & Berthelon (2009), por sua vez, avaliaram os efeitos do programa chileno de educação integral (um dos mais abrangentes da América Latina) na probabilidade de gravidez na adolescência e encontraram efeitos positivos do programa (aumento da educação integral diminui a probabilidade de gravidez na adolescência).

Este trabalho apresenta a primeira avaliação de impacto rigorosa do Programa Mais Educação em variáveis não relacionadas ao desempenho acadêmico. Avaliamos o PME em indicadores sociais (trabalho infantil e desemprego das mães) e indicadores relacionados à satisfação dos pais com a escola e ao comportamento dos estudantes (hábitos de leitura, hábitos de estudo, taxas de estudantes em recuperação, entre outros). Para isso, exploramos a descontinuidade identificada por Oliveira & Terra (2016) nos critérios de elegibilidade no programa a partir de 2012, quando o Ministério da Educação passou a privilegiar a escolha de escolas com mais de 50% dos alunos beneficiários do Programa Bolsa Família (PBF). A mensuração das variáveis de interesse foi realizada com base nos questionários preenchidos no âmbito do Sistema de Avaliação do Rendimento Escolas do Estado de São Paulo – Saresp. Nesse sentido, nossa análise se restringe às escolas situadas nesse estado.

Inicialmente, mostramos que a participação no programa aumenta em mais de 100% em escolas com um percentual de alunos beneficiários do PBF ligeiramente superior a 50% (grupo tratamento) em comparação com aquelas cujo percentual de estudantes beneficiários é ligeiramente inferior a 50% (grupo controle). Em seguida, usamos essa descontinuidade para avaliar o impacto do programa, comparando a média das variáveis de interesse em escolas tratadas e não-tratadas ao redor desse valor de referência (*cutoff*).

Adicionalmente, de forma a superar eventuais limitações de validade externa da metodologia de regressão descontínua, avaliamos o programa também pelo método de pareamento por escore de propensão com diferenças em diferenças. Para isso, selecionamos um conjunto de características observáveis e, a partir dessas características, buscamos encontrar, para cada escola no PME uma escola que não esteja no programa e que apresente características bastante semelhantes. Em seguida, aplicamos o método de diferenças-em-diferenças somente nas unidades pareadas.

Em linha com os trabalhos que avaliaram o impacto do PME em indicadores acadêmicos, não encontramos qualquer evidência de impacto significativo do programa nos indicadores sociais e de comportamento. Tanto pelo método de regressão descontínua como pelo método de diferenças em diferenças, a ampliação da jornada escolar não afetou significativamente e consistentemente nenhuma das variáveis analisadas.

O restante deste artigo está organizado em 9 seções. Inicialmente, na seção 2.2, apresentaremos as principais características do Programa Mais Educação e de sua implementação. A seção 2.3 apresentará as bases de dados e a seção 2.4 discutirá, em detalhes, a estratégia empírica e os modelos econométricos utilizados. Nas seções 2.5, e 2.6, analisaremos os principais resultados e realizaremos alguns testes de robustez. Já a seção 2.7 traz os resultados da estimação por diferenças-em-diferenças. Finalmente, a última seção discute os principais resultados do trabalho e as consequências em termos de políticas públicas.

2.2 O Programa Mais Educação

O Programa Mais Educação é uma política federal de transferência de recursos às escolas estaduais e municipais com vistas à implementação de atividades no contraturno escolar. Essas atividades têm por finalidade a melhoria do desempenho educacional, a redução da evasão, da reprovação e da distorção idade/série, entre outros

objetivos³⁰. O programa teve início em 2008 e, desde então, apresentou crescimento significativo no número de escolas participantes – de 1.408 escolas, no primeiro ano, para 58.652 escolas, em 2014.

Os recursos do PME são alocados à escola por meio do Programa Dinheiro Direto na Escola – PDDE/Educação Integral. Esses recursos devem ser utilizados para despesas de custeio e capital, tais como: o ressarcimento de despesas de transporte e alimentação de monitores e tutores responsáveis pelo desenvolvimento das atividades do PME³¹ e a aquisição dos materiais pedagógicos e de bens necessários às atividades, de acordo com os *kits* sugeridos.

Conforme previsto no Decreto nº 7.083, de 2010, que regulamenta o programa, diversas atividades podem ser desenvolvidas pelas escolas, tais como: acompanhamento pedagógico; experimentação e investigação científica; cultura e artes, esporte e lazer; cultura digital; educação econômica; comunicação e uso de mídias; meio ambiente; direitos humanos; prática de prevenção aos agravos à saúde; e promoção da saúde e da alimentação saudável. As escolas urbanas devem escolher de três a quatro macrocampos e quatro atividades a serem desenvolvidas com os alunos³². A atividade “Orientação de Estudos e Leitura”, do macrocampo Acompanhamento Pedagógico, é obrigatória e deve ser realizada com duração de uma hora a uma hora e meia diariamente³³.

Os estudantes inscritos no PME devem ter, pelo menos, sete horas diárias (ou 35 horas semanais) de atividades escolares. Cada turma do programa deve ser formada por 30 estudantes, exceto para atividades de Orientação de Estudos e Leituras e Campos

³⁰ O Decreto nº 7.083, de 2010, reforça, em seu art. 1º, esses objetivos, ao afirmar que o Programa Mais Educação tem por finalidade *contribuir para a melhoria da aprendizagem por meio da ampliação do tempo de permanência de crianças, adolescentes e jovens matriculados em escola pública, mediante oferta de educação básica em tempo integral*. No parágrafo primeiro deste artigo, o legislador define o que se entende por educação integral: *considera-se educação básica em tempo integral a jornada escolar com duração igual ou superior a sete horas diárias, durante todo o período letivo, compreendendo o tempo total em que o aluno permanece na escola ou em atividades escolares em outros espaços educacionais*.

³¹ Monitores são os indivíduos responsáveis pelas atividades do PME. Os tutores são responsáveis por atividades específicas e diferenciadas realizadas com alunos entre 15 e 17 anos que ainda estejam no ensino fundamental.

³² Esses critérios variaram pouco ao longo do tempo. A título de exemplo, em 2008, a quantidade mínima de atividades era três e era possível escolher apenas três macrocampos.

³³ As atividades relacionadas a cada macrocampo podem ser consultadas na Tabela A.1. do Apêndice.

do Conhecimento, cujas turmas deverão ser de 15 alunos. As turmas devem ser compostas por alunos de idades e séries variadas (de acordo com as características de cada atividade).

O Ministério da Educação recomenda que haja a participação de, no mínimo, 100 estudantes no PME por escola e que todos os estudantes da escola participem das atividades. Contudo, trata-se apenas de uma recomendação. Cada escola possui autonomia para, dentro de seu projeto político-pedagógico específico, definir quantos e quais alunos participarão das atividades.

Assim, o fato de uma escola integrar o Mais Educação não significa que todos os alunos dessa escola estão matriculados em educação integral. A participação de um aluno depende tanto de seu interesse como dos critérios utilizados pela escola no momento da adesão ao programa. Contudo, segundo dados do MEC e do Inep, a grande maioria dos alunos de uma escola adere ao programa: entre 2008 e 2011, a mediana de alunos participantes por escola foi de 82% e a média de 77% (Fundação Itaú, 2015).

A participação no programa depende da adesão voluntária das escolas que atendem aos critérios de elegibilidade. O Decreto nº 7.083, de 2010, confere ao Ministério da Educação a competência para definir, a cada ano, os critérios de priorização do PME, utilizando, entre outros, dados referentes i) à realidade da escola, ii) ao índice de desenvolvimento da educação básica e iii) às situações de vulnerabilidade social dos estudantes. Os critérios de adesão variaram ano a ano, conforme a Tabela 10. É importante ressaltar que esses são critérios de priorização, ou seja, escolas que não os atendem podem ter sua adesão ao programa autorizada. Trata-se de uma característica que afeta diretamente nossa estratégia empírica, como veremos a seguir.

Tabela 10 – Critérios de Elegibilidade do Programa Mais Educação

Critérios	2008	2009	2010	2011	2012	2013
Baixo IDEB	Menor que 3,5 e que a média municipal	Menor que 4,1	Sim (“baixo”)	Menor que 4,2 (séries iniciais) e 3,8 (séries finais)	Menor que 4,2 (séries iniciais) e 3,8 (séries finais)	Menor que 4,6 (séries iniciais) e 3,5 (séries finais)
População Municipal	Maior que 200 mil	Maior que 100 mil (regiões metropolitanas) 50 mil (demais))	Maior que 163 (90) mil (capitais e RM)	Maior que 18.844	Não	Não
Alunos	+99	+99	+99	+99	+99	+99
Compromisso Todos pela Educação e PDDE	Sim	Sim	Sim	Sim	Sim	Sim
Participação Anterior	Não	Sim	Sim	Sim	Sim	Só se IDEB < 3,5
Áreas de Vulnerabilidade Social	Sim	Sim	Sim	Plano Brasil Sem Miséria e Rurais	Plano Brasil Sem Miséria e Rurais	Rurais
Bolsa Família	Não	Não	Não	+ 50% dos alunos	+ 50% dos alunos	+ 50% dos alunos
Outros	-	-	-	Programa Escola Aberta	Programa Escola Aberta	Todos os Municípios

Fonte: Fundação Itaú Social e Grupo Banco Mundial. Programa Mais Educação: Avaliação de Impacto e Estudo Qualitativo (Apresentação dos Resultados)

Em 2011, teve início a parceria entre o MEC e o Ministério do Desenvolvimento Social – MDS, com o intuito de integrar o PME ao Programa Brasil sem Miséria (PBM). Com isso, buscava-se aprimorar a focalização do Programa, com maior ênfase a critérios relacionados à vulnerabilidade social dos estudantes e das famílias atendidas. Uma das mais importantes medidas nesse sentido foi a adoção do percentual de alunos oriundos de famílias beneficiárias do Programa Bolsa Família (PBF) como critério principal de elegibilidade das escolas. Criou-se, assim, o conceito de “escolas maioria PBF”: escolas nas quais pelo menos 50% dos alunos pertencem a famílias beneficiadas pelo Bolsa Família³⁴.

A partir de 2012, temos, assim, uma descontinuidade nos critérios de adesão ao programa: escolas maioria PBF passaram a ser priorizadas. Como veremos, essa descontinuidade não é determinística. É possível observar, nos anos seguintes, a participação no programa de escolas com menos de 50% de alunos beneficiários do Bolsa Família. Contudo, o limite (*cutoff*) de 50% marca uma mudança probabilística significativa: escolas à direita do *cutoff* passam a apresentar uma *probabilidade muito maior* de adesão ao Programa. Tem-se, assim, a possibilidade de avaliação de impacto do programa por meio de modelos de regressão descontínua *fuzzy*, que serão detalhados na seção 2.4.

2.3 Bases de Dados

Nesta seção, apresentaremos, de forma sucinta, as quatro bases de dados utilizadas em nossa análise.

A primeira base foi obtida junto ao Ministério da Educação e reúne as escolas participantes do Programa Mais Educação desde 2008. Trata-se da base utilizada para construção de nossa variável de tratamento.

A segunda base foi obtida junto ao Ministério do Desenvolvimento Social e reúne os alunos beneficiários do Programa Bolsa Família em cada escola. Utilizamos essa base para construção da nossa variável de elegibilidade. O acompanhamento dos alunos PBF é bimestral. Com base em documentos de assessoramento (Melo, 2015), adotamos o

³⁴ Os critérios foram adotados, principalmente, para escolas urbanas com mais de 100 alunos. Os critérios para escolas rurais foram, em geral, relacionados a características do ambiente em que as escolas estavam inseridas (ex. índices de pobreza do campo no município, índices de não alfabetização, etc.).

semestre de outubro/novembro como referência para calcular o total de alunos PBF da escola.

Outro ponto relevante a respeito da construção dessa variável diz respeito ao ano de referência. Iremos avaliar o programa em escolas que aderiram ao PME em 2012, primeiro ano afetado pelo critério de 50%. Nesse caso, os critérios de priorização e as escolas submetidas à tratamento foram definidos em 2011, razão pela qual adotamos o ano de 2010 como referência para o cálculo do percentual de alunos PBF.

Logo, a variável “percentual de alunos PBF” foi construída tendo por numerador o total de alunos PBF da escola em 2010 (medida em outubro/novembro desse ano) e como denominador o total de alunos matriculados no Ensino Básico da escola em 2010 (provenientes da base do Censo Escolar)³⁵. Essas variáveis não são preenchidas no mesmo momento do ano (o Censo é preenchido em maio), o que pode gerar imprecisões, resultantes da mobilidade ou evasão de alunos³⁶.

A terceira base reúne os dados do Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo – Saresp. Trata-se de avaliação externa conduzida, desde 1996, pela Secretaria da Educação do Estado de São Paulo. Inicialmente voltada para a rede estadual de ensino, atualmente a avaliação é aberta, por meio de adesão, também a escolas das redes municipais e particulares. O exame é aplicado no final de cada ano letivo para alunos do 3º, 5º, 7º e 9º anos do Ensino Fundamental³⁷ e 3º ano do Ensino Médio.

A base do Saresp de 2012 foi utilizada para mensurar os dados relativos ao trabalho infantil (no caso de estudantes do 9º ano) e à inserção da mãe no mercado de trabalho (para estudantes do 5º e 9º anos), além dos indicadores comportamentais (satisfação com a escola, hábitos de leitura e hábitos de estudo em casa) e acadêmicos (taxas de estudantes em recuperação). Para isso, utilizamos os questionários preenchidos pelos alunos e por seus pais. A base do Saresp foi utilizada, ainda, para mensuração de variáveis de controle (ex. perfil racial das famílias de cada escola e escolaridade dos pais).

³⁵ Como forma de endereçar a possibilidade de arredondamento na seleção das escolas para o Programa, utilizamos como referência para unidade acima do *cutoff* um percentual de alunos PBF maior ou igual a 49,5%.

³⁶ Em 2 escolas foi constatado um percentual superior a 100% de alunos PBF. Essas escolas foram desconsideradas na análise.

³⁷ A partir de 2013, a avaliação passou a considerar alunos do 2º ano do Ensino Fundamental.

Por fim, utilizamos os microdados do Censo Escolar de 2012 para mensuração de variáveis de controle relativas às características das escolas. Os microdados do Censo podem ser obtidos no site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. Além disso, no mesmo site, é possível obter os dados relativos às taxas de abandono, reprovação, aprovação, distorção idade-série e o Ideb³⁸ por escola. Esses dados também foram utilizados para a construção das variáveis de controle.

No Censo Escolar de 2012, identificamos cerca de 33 mil escolas no Estado de São Paulo. Após o tratamento³⁹ e cruzamento com a base de escolas participantes do Saresp, nossa base de Ensino Fundamental possui 6.586 escolas, sendo 4.016 na base dos Anos Iniciais e 3.640 na base de Anos Finais (1.070 escolas estão tanto na base dos Anos Iniciais quanto dos Anos Finais)⁴⁰. Dessas 6.586 escolas, 605 participaram do PME em 2012.

2.4 Estratégia Empírica

Como destacam Cattaneo & Vazquez-Bare (2016), a técnica conhecida como Regressão Descontínua – RDD é, provavelmente, a abordagem quantitativa mais reconhecida e internamente válida para análise e interpretação em um conjunto de dados não-experimentais. Isso porque requer hipóteses mais fracas de identificação do que outros métodos não-experimentais – como, por exemplo, diferenças em diferenças. Contudo, é também uma das mais exigentes em termos de informação, uma vez que exige que a probabilidade de receber o tratamento se altere de forma descontínua em alguma variável (no nosso caso, percentual de alunos PBF nas escolas).

Em termos intuitivos, nossa estratégia empírica consiste em i) avaliar se, de fato, há uma descontinuidade na participação no PME no *cutoff* (escolas com 50% de alunos PBF); ii) avaliar se essa descontinuidade é refletida em outras variáveis. A análise será

³⁸ O Índice de Desenvolvimento da Educação Básica (Ideb) é resultado da combinação de dois outros indicadores: i) pontuação média dos estudantes em exames padronizados ao final de cada etapa do ensino fundamental (5º e 9º anos) e do ensino médio (3º ano); ii) taxa de aprovação dos estudantes em cada etapa de ensino.

³⁹ Retiramos as escolas rurais, urbanas com menos de 100 alunos, privadas e públicas que não possuem Ensino Fundamental.

⁴⁰ Os Anos Iniciais correspondem ao período entre o 1º e o 5º ano do Ensino Fundamental; os Anos Finais ao período entre o 6º e o 9º ano.

conduzida no nível da escola. Para isso, agregamos os dados do Saesp (detalhados a nível do aluno) ao nível da escola, o que leva em consideração o fato de alunos da mesma escola não serem unidades independentes.

Seja Y_i a média de uma variável de interesse (y) para alunos matriculados na escola (i) - tal como a média dos alunos que trabalham na escola; X_i o percentual de alunos beneficiários do Programa Bolsa Família da escola (i) em 2010; e D o status do tratamento Assim, se $D_i = 1$, a escola i participou do PME em 2012 e se $D_i = 0$ a escola i não participou do programa em 2012.

Conforme demonstram Hahn, Todd, & Klaauw (2001), caso o tratamento seja definido de forma determinística (ou seja, todas escolas acima de 50% de alunos PBF são tratadas e todas abaixo desse limite não o são) e as funções de regressão $E[Y_i(0)|X_i = x]$ e $E[Y_i(1)|X_i = x]$ sejam contínuas em $x = \bar{x} = 50\%$, o efeito médio do tratamento no *cutoff* pode ser calculado por:

$$\tau_{SRD} = E[Y_i(1) - Y_i(0)|X_i = \bar{x}] = \lim_{x \downarrow \bar{x}} E(Y_i|X_i = x) - \lim_{x \uparrow \bar{x}} E(Y_i|X_i = x) \quad (1)$$

Assim definimos o efeito médio do tratamento em uma RDD *sharp*⁴¹ (τ_{SRD}). Trata-se do parâmetro que busca estimar o que teria ocorrido com a variável Y se as unidades controles com $X_i = \bar{x}$ tivessem aderido ao programa⁴².

No caso de um desenho *fuzzy*, o tratamento não é definido de forma determinística. Trata-se exatamente do caso do PME: por um lado, diversas escolas com maioria PBF não aderiram ao PME (como vimos na seção 2.2, a adesão é voluntária); por outro, há escolas com minoria PBF que tiveram sua adesão autorizada. Assim, o limite de 50% representa uma mudança na probabilidade, mas não de 0 para 1. A expectativa é que as escolas acima desse limite tenham uma probabilidade de adesão superior às escolas abaixo desse limite.

Vamos definir $T_i = 1$ se $x_i > 50\%$ e $T_i = 0$ se $x_i < 50\%$. Note que em um contexto *sharp*, $T_i = D_i$ para todo i . Contudo, em um desenho *fuzzy*, há algumas unidades para as

⁴¹ Por “*sharp*” entenda-se uma RDD em que a descontinuidade seja determinística: todos acima da descontinuidade são tratados e todos abaixo não o são.

⁴² Pela hipótese de continuidade, sabemos que é zero a probabilidade de que exista uma unidade i tal que $X_i = \bar{x}$ - trataremos desse problema a seguir.

quais $T_i \neq D_i$. Nesse caso, a participação no tratamento passa a ter dois valores potenciais $D_i(1)$, quando a unidade é assinalada para o grupo tratamento ($T_i=1$) e $D_i(0)$, quando a unidade é assinalada para o grupo controle ($T_i=0$).

A RDD *fuzzy* pode ser compreendida, assim, como uma estratégia de estimação baseada em variáveis instrumentais, na qual a descontinuidade se torna uma variável instrumental para o *status* do tratamento (ao invés de definir de forma determinística se a unidade é tratada ou não) (Lee & Lemieux, 2010) (Angrist & Pischke, 2009). Consequentemente, o parâmetro RDD *fuzzy* pode ser facilmente compreendido como uma razão entre duas regressões descontínuas *sharp*:

$$\tau_{FRD} = \frac{\lim_{x \downarrow \bar{x}} E(Y_i | X_i = x) - \lim_{x \uparrow \bar{x}} E(Y_i | X_i = x)}{\lim_{x \downarrow \bar{x}} E(D_i | X_i = x) - \lim_{x \uparrow \bar{x}} E(D_i | X_i = x)} \quad (2)$$

onde o denominador é o termo correspondente ao primeiro estágio (o efeito médio da elegibilidade para o tratamento sobre o tratamento), em um contexto de variáveis instrumentais. Já o numerador pode ser compreendido como a forma reduzida ou, ainda, como um estimador da “intenção de tratar” (ITT) (Cattaneo & Titiunik, 2017). Assim, a análise de uma regressão descontínua *fuzzy* é bastante semelhante à análise de uma *sharp*.

Note, ainda, que há duas formas de lidar com um contexto *fuzzy*. Podemos redefini-lo como uma RDD *sharp* e avaliar o impacto da “elegibilidade para o tratamento” nas variáveis de interesse. Nesse caso, avaliamos o impacto de ser elegível para o PME (e não de ter aderido ao programa) – ou o estimador da intenção de tratar (ITT). Ou podemos analisá-lo com base em uma estratégia de variáveis instrumentais, onde o parâmetro *fuzzy* pode ser compreendido como a razão entre dois parâmetros *sharp*. Nesse caso, avaliamos o impacto de ter aderido ao PME – ou o efeito local médio do tratamento (LATE).

Nesse trabalho, seguiremos uma estratégia de estimação por variáveis instrumentais, usando a elegibilidade para o tratamento como um instrumento para a efetiva participação no programa. Em outras palavras, adotamos uma estratégia de estimação de regressão descontínua *fuzzy* na qual o status do tratamento (D_i) é determinado, de forma probabilística, como uma função descontínua do percentual de alunos bolsa família na escola (Lee & Lemieux, 2010) (Angrist & Pischke, 2009). Logo, buscaremos estimar o LATE.

Vimos, acima, que uma das hipóteses mais importantes da nossa estratégia empírica é a premissa de continuidade ao redor do *cutoff*. Esperamos que a média de alunos que trabalham na escola ou a média de mães desempregadas sejam correlacionados com o percentual de alunos bolsa família. Sob a premissa de que o relacionamento entre as variáveis é contínuo ao redor do *cutoff*, qualquer salto na variável de interesse nessa região pode ser interpretado como uma evidência de impacto do tratamento. Assim, o parâmetro τ_{FRD} irá mensurar o efeito causal do PME sobre as variáveis de interesse (desemprego da mãe, trabalho infantil, desempenho acadêmico, entre outras) no *cutoff*.

Como já mencionamos, sob a premissa de continuidade, não esperamos que haja observações tais $X_i = \bar{x}$, o que torna necessário extrapolarmos para observações fora do *cutoff*. Nesse caso, devemos utilizar métodos polinomiais para modelar ou aproximar a função de regressão $E(Y_i|X_i = x)$ em cada lado do *cutoff* (separadamente). Dessa forma, uma decisão extremamente importante em uma estratégia RDD é como essa aproximação é realizada. A validade da estimação está diretamente relacionada à capacidade dos modelos polinomiais de proverem uma descrição adequada de $(Y_i|X_i = x)$ nos dois lados do *cutoff*.

Em geral, essa aproximação pode ser feita de duas formas: globais (onde todas as observações são utilizadas para estimação) e locais (onde somente as observações ao redor do *cutoff* são utilizadas). Abordagens globais foram bastante utilizadas nos estágios iniciais da literatura sobre regressão descontínua (Cattaneo & Vazquez-Bare, 2016). Nesse caso, busca-se aproximar as funções de regressão em cada lado do *cutoff* usando polinômios de graus mais elevados (3 ou 4) em todas as observações. Contudo, trabalhos recentes mostram que esses polinômios tendem a ser boas aproximações no geral, mas não em pontos de fronteira – o que os torna inadequados no contexto de uma RDD, estratégia de estimação baseada, justamente, na estimativa dos efeitos do tratamento em limites (Gelman & Imbens, 2014)⁴³.

Dessa forma, para reduzir a chance de estimativas incorretas na fronteira, tem se tornado cada vez mais comuns a utilização de abordagens locais, realizando a aproximação em uma vizinhança ao redor da descontinuidade $[\bar{x} - h, \bar{x} + h]$:

⁴³ Como mostram os autores, polinômios de graus mais elevados tendem a rejeitar a hipótese nula (ausência de efeito do tratamento) muito mais do que deveria ocorrer. Além disso, o peso atribuído a observações longe do *cutoff* pode ser muito alto e varia bastante conforme o grau utilizado para a aproximação. Os resultados estimados também se mostram muito sensíveis ao grau do polinômio.

$$\tau_{SRD} = \lim_{x \uparrow \bar{x}} E(Y_i | X_i = \bar{x} - h < x < \bar{x} + h) - \lim_{x \downarrow \bar{x}} E(Y_i | X_i = \bar{x} - h < x < \bar{x} + h) \quad (3)$$

Com a utilização de métodos polinomiais locais, focamos a aproximação das funções de regressão somente ao redor do *cutoff* (as demais observações, fora do intervalo definido, são desconsideradas na análise).

Para realizar esse procedimento, são necessárias algumas definições *a priori*: a função *kernel* (que assinalará pesos não negativos a cada observação com base na sua distância do *cutoff*); o grau do polinômio utilizado na aproximação; e, principalmente, a largura do intervalo (h). Feita essas definições, basta ajustar uma regressão de mínimos quadrados ponderados i) nas observações acima do *cutoff* (o intercepto estimado será a estimativa de $E[Y_i(1)|X_i = x]$) e ii) nas observações abaixo do *cutoff* (o intercepto estimado será a estimativa de $E[Y_i(0)|X_i = x]$). O estimador *sharp* será, então, a diferença entre esses dois interceptos (Cattaneo & Titiunik, 2017).

A escolha da função *kernel*, em geral, ocorre entre uma função triangular (que confere maior peso a observações próximas do *cutoff* e declina linearmente à medida que as observações se distanciam do limite) e uma função uniforme (que confere peso uniforme a observações dentro do *cutoff*)⁴⁴. Ambas as funções conferem peso zero a observações fora do intervalo escolhido $[\bar{x} - h, \bar{x} + h]$. No tocante ao grau do polinômio, como as estimações, em geral, ocorrem em intervalos pequenos ao redor do *cutoff*, recomenda-se a utilização de polinômios de menor grau (1 ou 2).

A escolha da largura da banda é, certamente, a questão mais importante em uma estratégia de estimação por RDD. Trata-se do intervalo ao redor do *cutoff* que será utilizado para ajustar o polinômio local que aproxima as funções de regressão. A escolha da largura do intervalo sofre do chamado “*trade-off* viés-variância”. Por um lado, quanto menor o intervalo, menor o erro ou viés da aproximação polinomial. Por outro lado, maior a variância dos coeficientes estimados, uma vez que menos observações estarão disponíveis.

Outra forma de analisar a escolha do intervalo é enxergá-lo como um *trade-off* entre eficiência e robustez. Alguns métodos serão mais “eficientes” sob determinadas

⁴⁴ Note que empregar a função *kernel* uniforme é equivalente a estimar uma regressão linear simples sem pesos dentro do intervalo.

hipóteses. Contudo, serão também mais vulneráveis à violação dessas hipóteses (ex. definição arbitrária dos intervalos). Por outro lado, alguns métodos são atualmente considerados mais robustos, ainda que menos precisos (ex. intervalos *data driven*). Infelizmente, não há uma determinação clara na literatura sobre o melhor método de escolha desse intervalo.

Em diversos trabalhos, a escolha da banda ocorre de forma *ad hoc* pelos pesquisadores, de forma a tentar balancear os objetivos de permanecer tão perto quanto possível do *cutoff*, ao mesmo tempo em que se conserva um número razoável de observações semelhantes entre si dentro do intervalo. Para isso, a estratégia é, geralmente, apresentar as estimativas para um intervalo amplo de candidatos a largura de bandas.

Neste trabalho, adotaremos três intervalos ao redor do *cutoff*: 5 pontos percentuais (ou seja, serão consideradas escolas com 45% a 55% alunos PBF), 10 pontos percentuais (escolas com 40% a 60% de alunos PBF) e 15 pontos percentuais (escolas com 35% a 65% de alunos PBF). Nesse caso, vamos estimar regressões locais lineares para estimar os limites a direita e a esquerda da descontinuidade usando pesos *kernel* uniformes. O impacto estimado será dado pela diferença entre os dois limites. Como veremos na seção 2.6, o intervalo de 10 é o que melhor parece atingir o objetivo de permanecer tão perto quanto possível do *cutoff* e, ao mesmo tempo, contar com um número razoável de observações estatisticamente semelhantes.

Como destacam Cattaneo & Vazquez-Bare (2016), uma importante premissa por trás desse procedimento é que a função de regressão está corretamente especificada ou, pelo menos, que os erros de má especificação são pequenos o suficiente para serem ignorados. Tal premissa é fundamental para a utilização de estatísticas de mínimos quadrados ordinários para estimação e inferência. Em outras palavras, vamos assumir que a função de regressão linear local é capaz de aproximar corretamente a distribuição real, uma premissa que, naturalmente, é mais plausível quanto menor o intervalo adotado. No *trade-off* eficiência vs. robustez, a definição *ad hoc* dos intervalos é um método eficiente, desde que se assuma que a função de regressão está corretamente especificada.

Como as estimativas são, muitas vezes, altamente sensíveis à escolha das bandas, a literatura mais recente tem optado por também apresentar os resultados para métodos de seleção orientados pelos dados (*data-driven*), de forma a se contornar o problema de conferir excessiva discricionariedade ao pesquisador (o que ocorre no caso de bandas definidas de forma *ad hoc*) e conferir maior robustez às estimativas. Esses métodos de seleção buscam balancear o *trade-off* viés-variância de uma forma ótima.

O método mais adotado atualmente foi desenvolvido por Imbens & Kalyanaraman (2012) e busca minimizar o erro mínimo quadrado do estimador RDD, dada a escolha da ordem do polinômio e a função *kernel*, razão pela qual é conhecido como MSE-ótimo⁴⁵. Como o erro quadrado mínimo do estimador é uma composição do quadrado do viés e da variância, o que o método faz é tentar “otimizar” o *trade-off* viés-variância. A descrição do processo de seleção da largura do intervalo MSE-ótimo pode ser encontrada em detalhes em Imbens & Kalyanaraman (2012) e Calonico et al. (2014).

O método permite, ainda, a escolha de bandas ótimas diferentes para o grupo tratamento e o grupo controle, o que pode ser interessante no caso em que o viés e/ou variância dos grupos controles e tratamento diverjam substancialmente (o que pode acontecer, por exemplo, no caso de diferenças significativas no número de observações em cada lado do *cutoff*, como é o caso da nossa variável de elegibilidade). Tanto com bandas uniformes quanto com bandas diferentes nos dois lados do *cutoff*, teremos um estimador RDD MSE-ótimo.

Cattaneo & Vazquez-Bare (2016) ressaltam que o cálculo do intervalo MSE-ótimo deve ser realizado de acordo com o objetivo e as premissas estabelecidas. Assim, não se deve empregar o mesmo intervalo para diferentes variáveis de interesse ou para as covariadas. Em outras palavras, ao contrário do método de definição arbitrária de intervalos, o método MSE-ótimo exige que para cada variável de interesse seja calculado um intervalo ótimo distinto⁴⁶.

Adicionalmente, o estimador MSE-ótimo acima descrito não pode ser usado diretamente para inferência. Trata-se de uma conclusão lógica uma vez que, se o estimador é construído para minimizar o *trade-off* viés-variância, ele é construído sob a premissa de que existe um viés, o que torna a inferência inválida. Por essa razão, na análise desse estimador, adotaremos a abordagem de inferência desenvolvida por Calonico et al. (2014), denominada correção robusta do viés.

⁴⁵ O método foi posteriormente aperfeiçoado por Calonico et al. (2014).

⁴⁶ Conforme recomendam Cattaneo & Vazquez-Bare (2016), *do not employ the same neighborhood for different outcome variables, pre-intervention covariates (if conducting falsification testing), estimation and inference procedures, or falsification methods. Using the same neighborhood for different goals, outcomes or samples disregards the specific empirical features (e.g., number of observations near the cutoff, variability or curvature), and will lead to unreliable empirical results due to invalidity of the methods employed.*

Por fim, há basicamente duas técnicas de falsificação/validação usualmente aplicadas na literatura. A primeira visa testar a continuidade da densidade da variável de elegibilidade ao redor do *cutoff*. Com isso, busca-se descartar a existência de manipulação no valor dessa variável. Trata-se de um cenário pouco provável neste trabalho, uma vez que adotamos como variável índice o percentual de alunos PBF em 2010, um ano antes da criação do critério de elegibilidade. Ainda assim, conduzimos teste de densidade de forma a descartar a possibilidade de manipulação.

A segunda técnica é baseada em uma premissa importante da nossa estratégia de identificação: a “suavidade” da evolução das variáveis de interesse ao redor do *cutoff* na ausência do tratamento. Uma forma de corroborar essa premissa é testar se há evidência de efeito do tratamento (descontinuidade) em variáveis que não deveriam ser por ele afetadas. Para isso, utilizaremos os mesmos métodos aplicados para as variáveis de interesse em variáveis que não deveriam ser por ele afetadas (ex. perfil racial das famílias, escolaridade dos pais, infraestrutura das escolas, etc.).

Adicionalmente, no caso dos intervalos definidos arbitrariamente, conduzimos um teste de diferença de médias no grupo de escolas antes e depois do *cutoff*, de forma a verificar se estamos lidando com grupos controle e tratamento relativamente homogêneos.

2.5 Análise dos Resultados

2.5.1. Participação no Programa Mais Educação

No primeiro estágio, buscaremos avaliar se o critério de elegibilidade implementado, formalmente, em 2011, de fato se refletiu em uma maior probabilidade de seleção de escolas com maioria PBF em 2012.

Os resultados do primeiro estágio são calculados para três bases distintas: Ensino Fundamental (que reúne todas as escolas de Ensino Fundamental do Estado de São Paulo que participaram do Saresp em 2012); Anos Iniciais (que reúne somente as escolas com alunos matriculados do 1º ao 5º anos); e Anos Finais (que reúne somente as escolas com alunos matriculados do 6º ao 9º anos). Essa divisão é importante, uma vez que, intuitivamente, o impacto da educação integral no desemprego das mães é mais provável na base de Anos Iniciais (com estudantes mais jovens, de 7 a 11 anos). Já o impacto no trabalho infantil somente pode ser verificado na base dos anos finais (com estudantes mais velhos, de 12 a 15 anos), uma vez que essa pergunta não está presente no questionário para alunos dos Anos Iniciais.

Os resultados podem ser consultados na Tabela 11. Cada painel apresenta o cálculo para uma base de dados (Ensino Fundamental, Anos Iniciais e Anos Finais). Cada coluna apresenta os resultados de uma especificação. As colunas intituladas “intervalos *ad hoc*” apresentam os resultados em intervalos pré-definidos ao redor do *cutoff*: 5, 10 e 15. Já as colunas denominadas “intervalos *data driven*” traz os resultados das estimações com base em intervalos calculados de forma MSE-ótima - tanto para bandas uniformes quanto para bandas diferentes dos dois lados do *cutoff* - e com aproximação com polinômios de grau 1 e 2⁴⁷.

Para cada estimacão, apresentamos o estimador RDD do efeito do tratamento (diferença entre tratados e controles na probabilidade de adesão ao programa), o p-valor calculado de forma tradicional para os intervalos *ad hoc* e o p-valor robusto a viés para os intervalos data-driven. Apresentamos, ainda, a dimensão do intervalo ao redor do *cutoff* e o número de observações com peso diferente de zero à esquerda e à direita do *cutoff*. Por fim, a coluna “média dos controles” traz o limite da aproximação pela esquerda do *cutoff* (ou seja, a aproximação pelo grupo controle) utilizando a estimacão com um intervalo de 10. Essa média dá a dimensão do impacto do critério de elegibilidade na participação no PME.

Como se nota, o critério de elegibilidade implicou uma mudança significativa na probabilidade de participação no PME ao redor do *cutoff* nas três amostras. Para a amostra do Ensino Fundamental, observamos na coluna “média dos controles” que à esquerda do *cutoff*, apenas 6% das escolas aderiram ao PME em 2012, enquanto, à direita do *cutoff*, mais de 24% das escolas aderiram ao programa nesse ano - o que representa um aumento estimado na probabilidade de participação de 18p.p. Para os Anos Iniciais e Finais, o cenário é semelhante: um aumento de 13 p.p. no primeiro caso (de 7% para 20%) e de 21 p. p. no segundo caso (de 4% para 25%).

⁴⁷ Nas estimacões *data driven*, utilizamos, ainda, duas covariadas: alunos matriculados em 2010 e uma *dummy* que indica se a escola é estadual ou municipal.

Tabela 11 – Estimativas dos Efeitos da Elegibilidade para o PME na Participação no Programa em 2012 (primeiro estágio)

	Média dos Controles	Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i> (<i>polinômio de grau 1</i>)		Intervalos <i>Data-Driven</i> (<i>polinômio de grau 2</i>)	
		5	10	15	Bandas Uniformes	Bandas Diferentes	Bandas Uniformes	Bandas Diferentes
Painel A: Ensino Fundamental								
Coefficiente	0.06	0.13	0.18***	0.17***	0.19***	0.19***	0.19**	0.20***
p-valor		0.13	0.00	0.00	0.00	0.00	0.02	0.01
p-valor robusto		-	-	-	0.01	0.00	0.05	0.01
Intervalo		-	-	-	15	27 / 14	14	30 / 13
Observações à esquerda/direita		158 / 107	474 / 161	974 / 188	933 / 187	2660 / 181	811 / 179	3213 / 179
Painel B: Anos Iniciais								
Coefficiente	0.07	0.07	0.13**	0.12**	0.13*	0.14**	0.13	0.15*
p-valor		0.53	0.05	0.03	0.04	0.02	0.17	0.06
p-valor robusto		-	-	-	0.09	0.05	0.26	0.09
Intervalo		-	-	-	14	25 / 14	14	26 / 14
Observações à esquerda/direita		128 / 83	365 / 126	725 / 147	662 / 143	1633 / 142	622 / 140	1756 / 140
Painel C: Anos Finais								
Coefficiente	0.04	0.18	0.21*	0.21**	0.25**	0.24**	0.25*	0.26**
p-valor		0.21	0.06	0.02	0.03	0.02	0.05	0.02
p-valor robusto		-	-	-	0.05	0.02	0.08	0.02
Intervalo		-	-	-	9	19 / 8	13	28 / 12
Observações à esquerda/direita		66 / 36	206 / 52	422 / 61	165 / 51	667 / 48	308 / 59	1419 / 56

Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%.

Chama a atenção a consistência dos coeficientes estimados com relação aos intervalos. Na amostra dos Anos Iniciais, houve maior variação, com a magnitude do coeficiente sendo inferior no menor intervalo (5), o que pode causar pequena distorção nas estimações do segundo estágio realizadas nas seções seguintes⁴⁸. Contudo, o estimador se mostrou bastante coerente nos demais intervalos, assim como nos intervalos das estimações dos Anos Finais e do Ensino Fundamental.

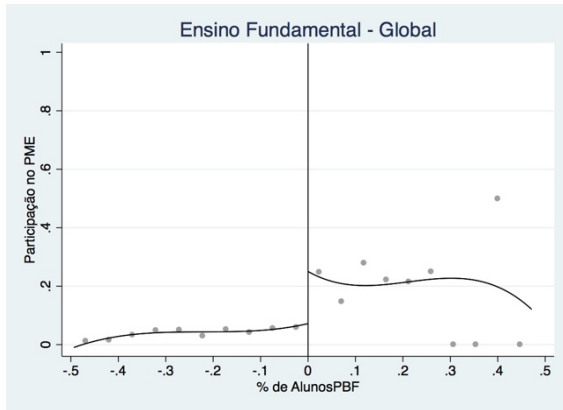
Dentre os intervalos definidos de forma *ad hoc*, o de 10 mostra-se o mais adequado. Isso porque, ao mesmo tempo em que conserva um número razoável de escolas tanto à direita quanto à esquerda do *cutoff*, essas unidades permanecem estatisticamente semelhantes entre si, o que não ocorre com o intervalo de 15 (vide seção 2.6). Adicionalmente, como era de se esperar, o p-valor aumenta significativamente no menor intervalo, em função do baixo número de observações, o que dificulta a estimação e a inferência dentro do intervalo de 5.

Os resultados estimados para o primeiro estágio podem ser observados graficamente na Figura 6. Os painéis (a), (b) e (c) ilustram uma aproximação global para as três bases, com o único objetivo de permitir uma visualização da tendência de toda a série. Já os painéis (d), (e) e (f) ilustram as estimativas locais lineares tendo por base a estimação com o intervalo de 10 (os gráficos com aproximações utilizando polinômios de grau 2 podem ser conferidos na Figura A.1 do Apêndice). Os pontos mostram médias das células a partir do agrupamento dos dados em cada lado do *cutoff*. Em todos os gráficos do artigo, centramos o *cutoff* em 0 – assim, escolas à direita do *cutoff* possuem mais de 50% dos alunos PBF e escolas à esquerda do *cutoff* possuem menos de 50% dos alunos PBF.

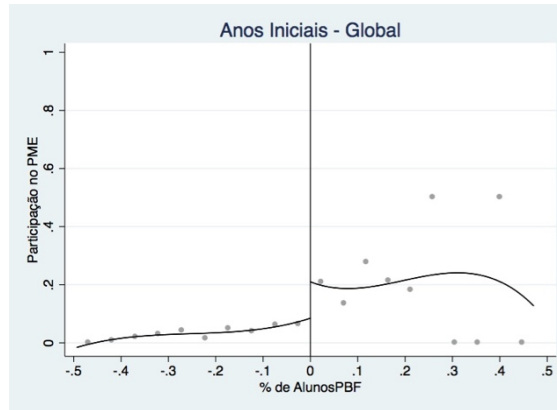
A análise gráfica confirma os resultados da Tabela 11. Como se nota, a descontinuidade no critério de elegibilidade adotado em 2012, de fato, se reflete em um aumento significativo na probabilidade de adesão ao Programa.

⁴⁸ Como vimos na equação (2), esse coeficiente é o denominador da estimação do “segundo estágio”, distorcendo, assim, os coeficientes estimados.

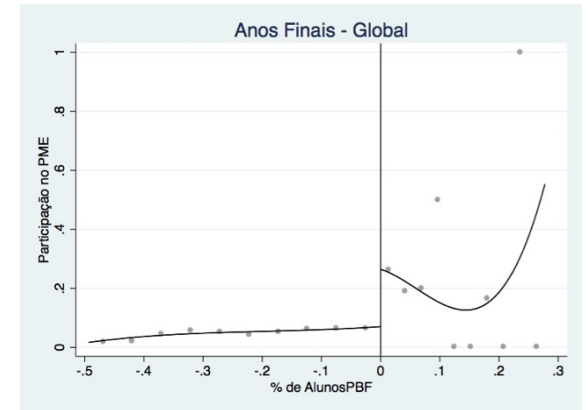
Figura 6 – Descontinuidade na Participação no PME em 2012



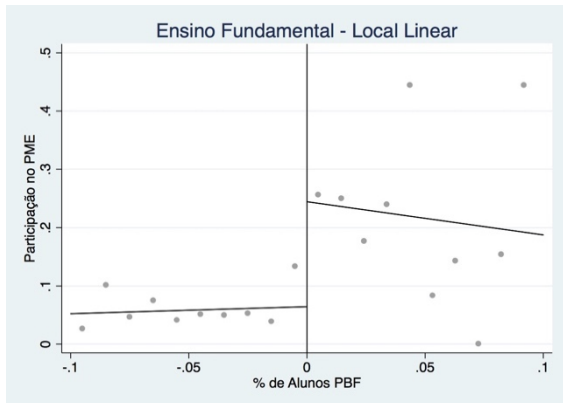
(a) Ensino Fundamental (global)



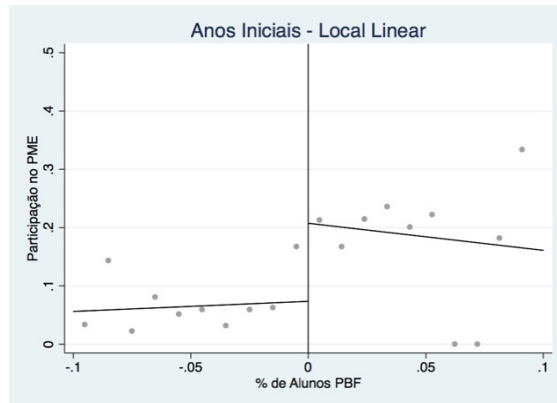
(b) Anos Iniciais (global)



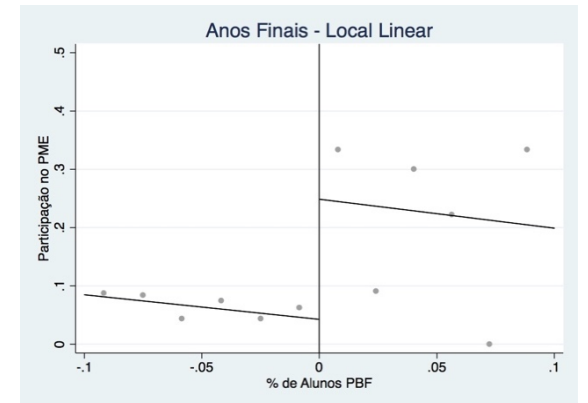
(c) Anos Finais (global)



(d) Ensino Fundamental (local linear)



(e) Anos Iniciais (local linear)



(f) Anos Finais (local linear)

Como forma de reforçar a robustez dos resultados estimados, estimamos os mesmos modelos para a participação no Programa Mais Educação em 2011, quando o critério de elegibilidade ainda não tinha sido adotado. Como se nota tanto pela Análise da Tabela A.2 quanto da Figura A.2 do Apêndice em nenhum dos casos a descontinuidade estimada mostrou-se estatisticamente ou economicamente significativa.

Nas seções seguintes, avaliamos o impacto do Programa Mais Educação nas variáveis de interesse. Para isso, estimamos uma versão da equação (2) para cada uma das variáveis, de forma a avaliar se a descontinuidade existente na probabilidade de participação no PME, que observamos na seção 5, reflete-se nas demais variáveis.

2.5.2. Programa Mais Educação e o Desemprego Materno

Para estimar o impacto do PME no desemprego das mães, lançamos mão dos questionários preenchidos na avaliação do Saesp. O questionário oferece 6 opções para a situação de trabalho da mãe: a) empregada; b) autônoma (trabalha por conta própria); c) dona de negócio próprio; d) trabalhadora temporária; e) aposentada; f) desempregada; e g) outra situação⁴⁹. Avaliamos o impacto em cada uma das respostas separadamente. Reportamos na Tabela 12 os resultados para “empregada”, “desempregada”, “trabalhadora temporária”, “autônoma” e “dona de negócio próprio”.

Como a maioria de nossas tabelas apresentará formato similar, descrevemos a Tabela 12 com maiores detalhes. Há duas grandes colunas - Anos Iniciais e Anos Finais - que trazem os impactos estimados para cada uma das bases. Apresentamos os resultados para cinco especificações distintas: três com base em “intervalos *ad hoc*” ao redor do *cutoff* (5, 10 e 15) e duas com base em intervalos *data driven* (com bandas uniformes e com bandas diferentes). Cada painel apresenta o impacto do programa em uma variável de interesse diferente. Em cada painel é informado o coeficiente estimado, p-valor tradicional e robusto (no caso dos intervalos *data driven*) e o intervalo no qual a

⁴⁹ Uma questão que poderia surgir é se o programa afeta não necessariamente o desemprego, mas a entrada das mães no mercado de trabalho. Buscamos analisar essa possibilidade avaliando o impacto da mãe na resposta “outra situação”, uma vez que é a única opção do questionário para mães que não trabalham e não buscam emprego. Nesse caso, esperaríamos um impacto negativo do programa (diminui a quantidade de mães fora do mercado), contudo, não encontramos impacto significativo do programa.

estimação foi calculada⁵⁰. O coeficiente foi estimado ao nível da escola e corresponde ao valor médio da variável para os estudantes/pais da escola. Assim, a título de exemplo, o coeficiente da taxa de desemprego deve ser interpretado como “o percentual de mães que se declararam desempregadas na escola i ”.

A análise da Tabela 12 mostra que o programa não impactou significativamente o desemprego das mães. Tanto para os Anos Iniciais quanto para os Anos Finais, não se observou qualquer variação significativa em nenhuma das 5 categorias de oferta de trabalho analisadas.

Esse resultado é ilustrado na Figura 7, que traz os resultados do primeiro estágio da aproximação local de um polinômio de grau 1 para as variáveis “mãe empregada” e “mãe desempregada” nos Anos Iniciais e Finais. Assim como na Tabela, também nos gráficos não observamos qualquer indicativo de descontinuidade significativa na evolução dessas variáveis ao redor do *cutoff*, o que demonstra a ausência de impacto do programa nessas variáveis.

⁵⁰ Note que no caso dos intervalos *ad hoc*, o intervalo é sempre igual para todas as variáveis. Já no caso de intervalos *data driven*, esse intervalo pode e tende a variar, conforme explicamos na seção 2.4.

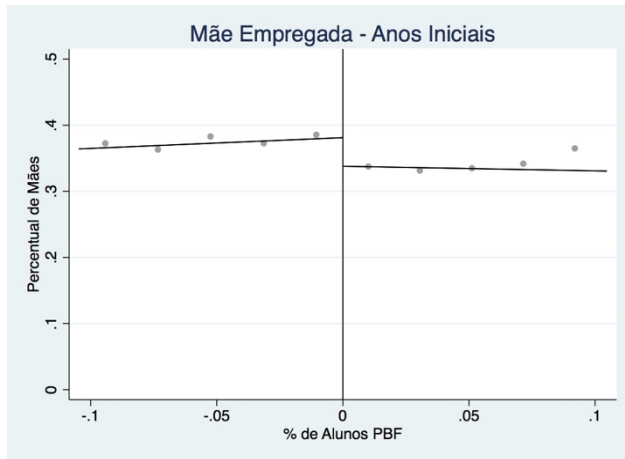
Tabela 12 – Efeitos do PME no desemprego da mãe (segundo estágio)

	Anos Iniciais					Anos Finais				
	Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Drive</i>		Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Drive</i>	
	5	10	15	Bandas Uniformes	Bandas Diferentes	5	10	15	Bandas Uniformes	Bandas Diferentes
Painel A: Mãe Empregada										
Coefficiente	-0.71	-0.33	-0.25	-0.28	-0.20	-0.27	-0.21	-0.10	-0.12	-0.10
p-valor	0.55	0.11	0.13	0.13	0.15	0.37	0.21	0.39	0.31	0.42
p-valor robusto				0.11	15				36	0.44
Intervalo				15	22 / 15				15	17 / 7
Painel B: Mãe Desempregada										
Coefficiente	0.31	0.22	0.13	0.19	0.16	0.07	0.11	0.13	0.05	0.05
p-valor	0.58	0.14	0.25	0.15	0.16	0.64	0.30	0.17	0.48	0.53
p-valor robusto				0.16	14				0.66	0.66
Intervalo				12	17 / 13				12	18 / 8
Painel C: Mãe Trabalhadora Temporária										
Coefficiente	-0.02	0.03	0.02	0.01	0.00	-0.10	-0.09	-0.07	-0.07	-0.08
p-valor	0.90	0.68	0.82	0.90	0.99	0.36	0.19	0.14	0.12	0.14
p-valor robusto				0.83	0.88				0.20	0.26
Intervalo				15	25 / 10				12	20 / 5
Painel D: Mãe Autônoma										
Coefficiente	0.05	0.08	0.07	0.06	0.06	0.04	0.02	-0.04	0.01	0.01
p-valor	0.82	0.25	0.32	0.36	0.35	0.69	0.80	0.51	0.90	0.86
p-valor robusto				0.59	0.75				0.94	0.89
Intervalo				15	24 / 9				13	16 / 07
Painel E: Mãe Negócio Próprio										
Coefficiente	-0.04	-0.01	-0.02	-0.01	0.01	0.02	0.01	-0.01	0.00	-0.01
p-valor	0.74	0.84	0.47	0.79	0.63	0.68	0.69	0.69	0.98	0.69
p-valor robusto				0.82	0.60				1.00	0.49
Intervalo				15	22 / 9				14	26 / 7

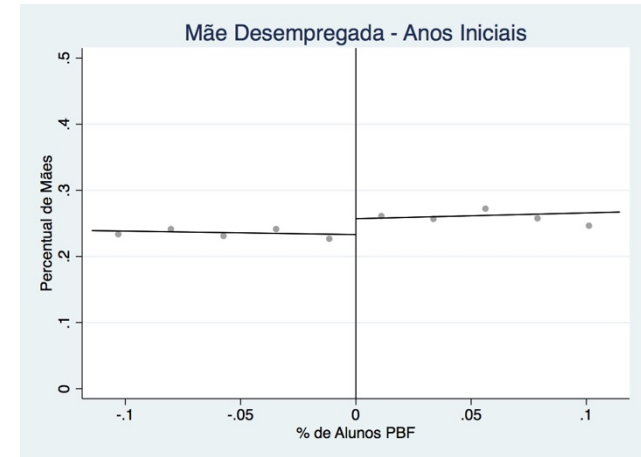
Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%.

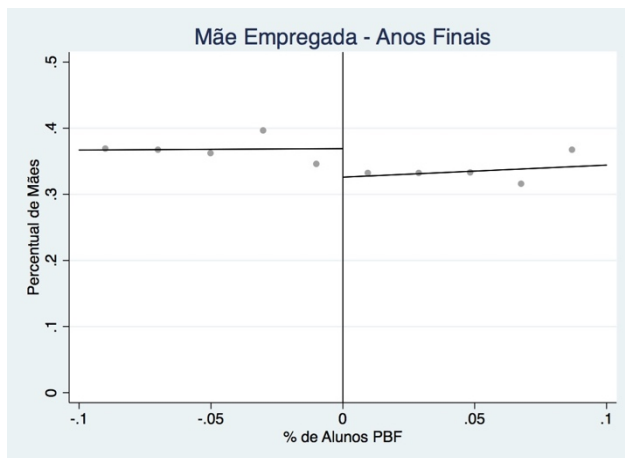
Figura 7 - Efeitos do PME no desemprego da mãe



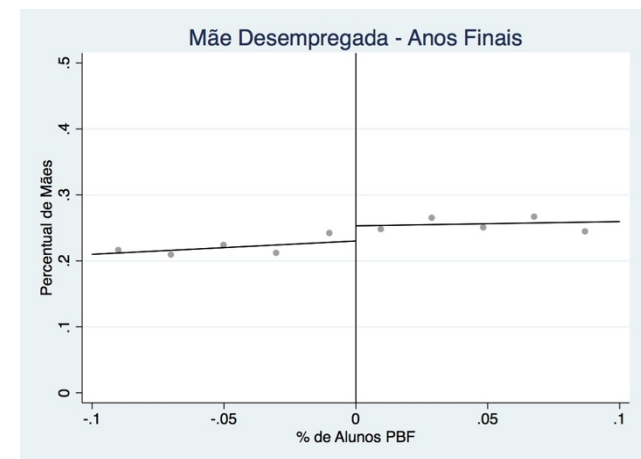
(a) Mãe empregada (Anos Iniciais)



(b) Mãe desempregada (Anos Iniciais)



(c) Mãe empregada (Anos Finais)



(d) Mãe desempregada (Anos Finais)

2.5.3. Programa Mais Educação e o Trabalho Infantil

Para estimar o impacto do PME no trabalho infantil, utilizamos somente a base de Anos Finais do Saresp (somente é feita essa pergunta para os alunos do 6º ao 9º anos).

A pergunta no questionário dos Anos Finais é sobre se o estudante trabalha em dias de aula. São oferecidas três possibilidades de respostas: a) não trabalho, só estudo; b) sim, trabalho em casa, ajudando nas tarefas de casa; e c) sim, trabalho fora de casa. As estimativas levam em consideração todas as respostas.

Reportamos na Tabela 13 os resultados para as respostas “b” (percentual de alunos que trabalham em casa por escola) e “c” (percentual de alunos que trabalham fora de casa por escola). Novamente, não há qualquer evidência de impacto significativo do PME. Esses resultados estão ilustrados nos painéis *a* e *b* da Figura 8.

Trata-se de uma conclusão surpreendente, uma vez que os alunos das escolas do programa devem permanecer pelo menos 7 horas por dia na escola, o que, ao menos em tese, os deixaria menos disponíveis tanto para o trabalho em casa, ajudando nas tarefas domésticas, como para o trabalho fora de casa. Não é o que se observa, contudo, na análise dos dados. Uma hipótese é que os alunos que trabalham não estejam aderindo ao programa. Outra hipótese é que haja simplesmente um deslocamento de horário no trabalho – ex. os estudantes passam a ajudar nas tarefas domésticas à noite. De qualquer forma, chama a atenção o fato de que a educação integral não tenha qualquer impacto no trabalho infantil, principalmente fora de casa.

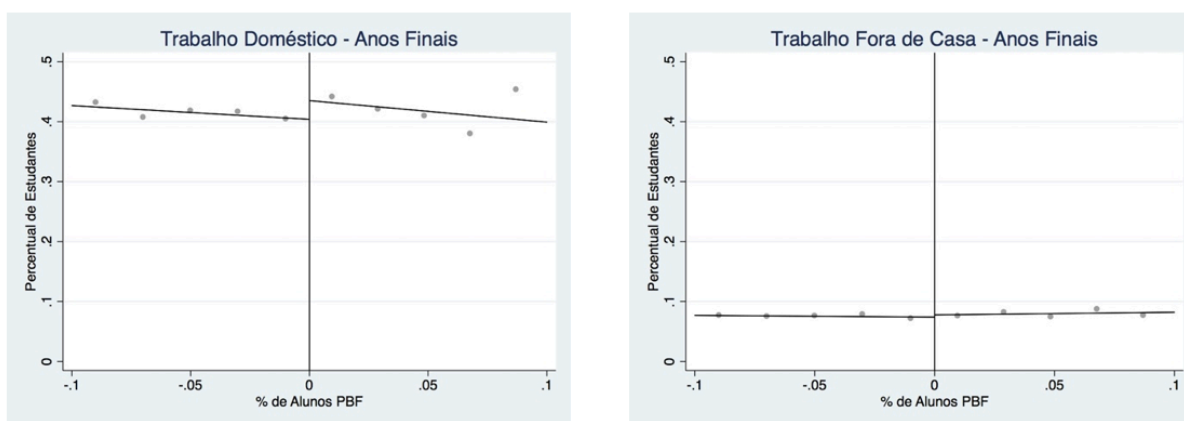
Tabela 13 – Efeitos do PME no trabalho infantil (segundo estágio)

	Anos Finais				
	Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Drive</i>	
	5	10	15	Bandas Uniformes	Bandas Diferentes
Painel A: Trabalha em casa					
Coefficiente	0.23	0.15	-0.01	0.06	0.07
p-valor	0.39	0.30	0.90	0.48	0.43
p-valor robusto	-	-	-	0.27	0.37
Intervalo	-	-	-	13	19 / 8
Painel B: Trabalha fora de casa					
Coefficiente	0.11	0.02	0.01	0.03	0.04
p-valor	0.30	0.78	0.78	0.45	0.43
p-valor robusto	-	-	-	0.41	0.33
Intervalo	-	-	-	17	21 / 6

Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%.

Figura 8 – Efeitos do Programa Mais Educação no trabalho infantil



(a) Trabalha em casa (Anos Finais)

(b) Trabalha fora de casa (Anos Finais)

2.5.4. Programa Mais Educação e a Satisfação dos Pais com a Escola

A implementação de programas de educação integral pode ter efeito, ainda, na satisfação dos pais com a escola e na percepção dos pais sobre o interesse dos filhos pela educação. Avaliamos essa hipótese por meio de quatro perguntas do questionário, nas quais se pede para que os pais deem uma nota de zero a 10 para os seguintes itens: i) interesse dos filhos pelos estudos, ii) professores da escola; iii) infraestrutura da escola (instalações físicas); e iv) segurança da escola.

Como se nota na Tabela 14, o programa não impactou significativamente a percepção dos pais do interesse dos filhos pelos estudos nem a imagem que possuem dos professores da escola – tanto para os anos iniciais quanto para os anos finais.

Há, contudo, um indicativo de impacto positivo do PME na nota que os pais conferem à infraestrutura da escola para a amostra dos Anos Finais, sendo o coeficiente significativo a 10% no intervalo de 10 e a 5% no intervalo de 15 e nos intervalos ótimos. Esse impacto está ilustrado no Painel (e) da Figura 9 e pode decorrer do fato de o programa ter gerado melhoras na infraestrutura das escolas, por meio da construção de espaços para atender as atividades do contraturno⁵¹. Outra variável em que se nota pequeno indicativo de impacto positivo é na nota conferida à segurança da escola – vide painel (f) da Figura 9. No entanto, nesse caso, o impacto observado é somente ao nível de 10% e nas estimações com intervalos ótimos.

Em ambos os casos, o impacto não é verificado na amostra dos Anos Iniciais e nem na estimação por diferenças-em-diferenças (vide seção 2.7).

⁵¹ Ao longo dos anos, o Ministério da Educação previu repasses adicionais de recursos para atividades relacionadas ao programa. A título ilustrativo, Resolução nº 62, de 14 de dezembro de 2009, autorizava a transferência de recursos para escolas públicas do Programa Mais Educação destinados a reforma, ampliação e construção de cobertura de quadras esportivas ou de espaços destinados ao esporte e ao lazer.

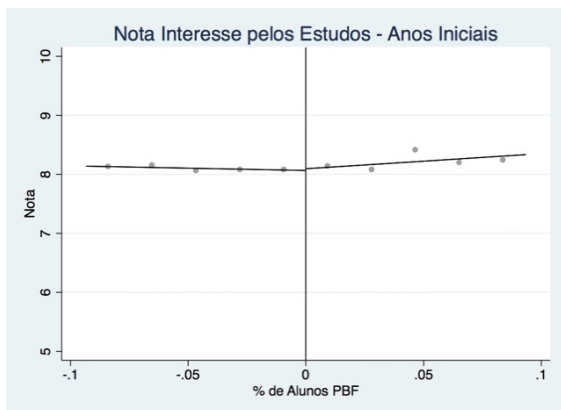
Tabela 14 – Efeitos do PME na satisfação dos pais com a escola (segundo estágio)

	Anos Iniciais					Anos Finais				
	Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i>		Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i>	
	5	10	15	Bandas Uniformes	Bandas Diferentes	5	10	15	Bandas Uniformes	Bandas Diferentes
Painel A: Nota Interesse Estudos (Pais)										
Coeficiente	0.41	0.56	0.34	0.18	-0.12	0.63	0.42	0.22	0.21	0.18
p-valor	0.86	0.42	0.54	0.76	0.80	0.53	0.49	0.62	0.66	0.63
p-valor robusto				0.70	0.88				0.65	0.47
Intervalo				15	22 / 10				0.13	22 / 6
Painel B: Nota dos Professores										
Coeficiente	-2.45	-0.74	-0.59	-1.05	-1.15	1.07	0.96	0.75	0.62	0.77
p-valor	0.46	0.27	0.31	0.13	0.08	0.38	0.19	0.16	0.20	0.14
p-valor robusto				0.10	0.09				0.26	0.14
Intervalo				15	25 / 9				0.14	20 / 8
Painel C: Nota Infraestrutura										
Coeficiente	7.96	1.94	1.57	1.47	0.93	4.25	4.49*	3.12**	3.58**	3.90**
p-valor	0.56	0.25	0.29	0.29	0.38	0.25	0.08	0.05	0.03	0.02
p-valor robusto				0.24	0.33				0.04	0.02
Intervalo				15	20 / 12				0.12	19 / 7
Painel D: Nota da Segurança										
Coeficiente	5.39	1.86	1.60	0.82	0.39	2.97	3.51	2.55	2.44*	3.33*
p-valor	0.60	0.30	0.31	0.52	0.69	0.35	0.12	0.11	0.08	0.05
p-valor robusto				0.55	0.78				0.10	0.07
Intervalo				15	22 / 9				16	19 / 7

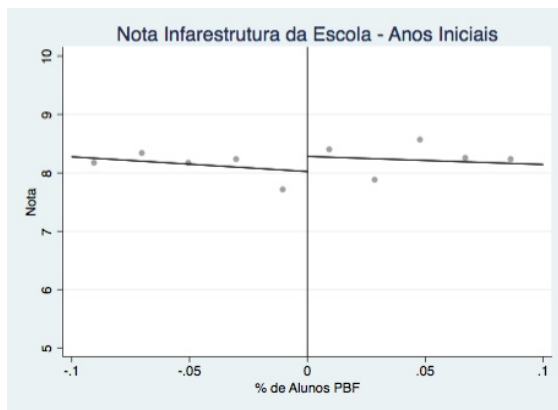
Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%.

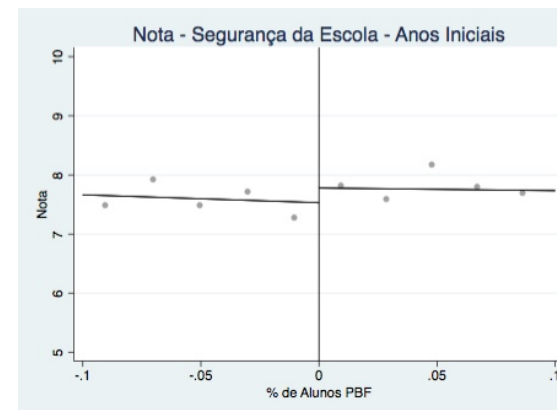
Figura 9 - Efeitos do PME na Satisfação dos Pais com a Escola



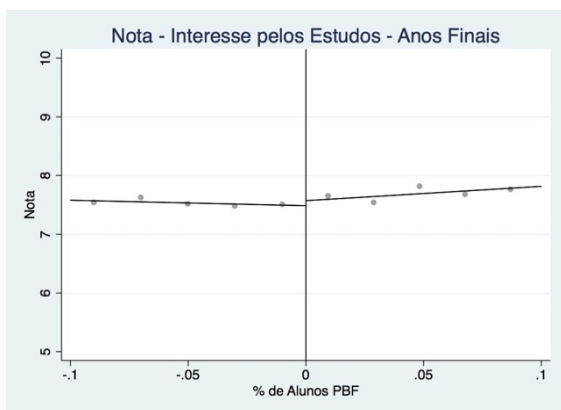
(a) Nota – Interesse pelos Estudos (Anos Iniciais)



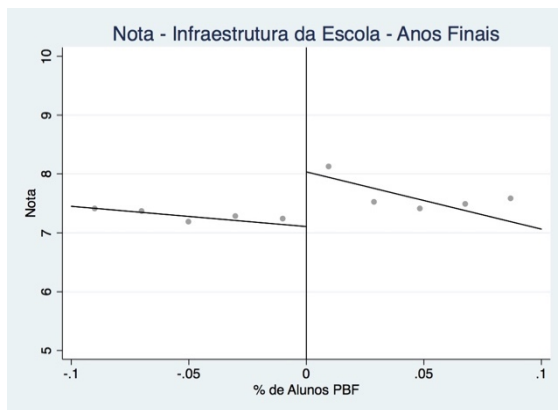
(b) Nota – Infraestrutura da Escola (Anos Iniciais)



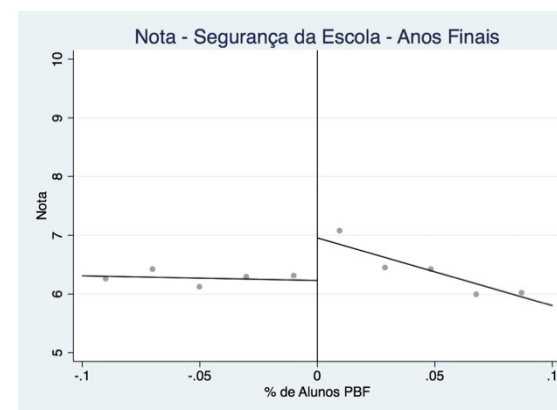
(c) Nota – Segurança da Escola (Anos Iniciais)



(d) Nota – Interesse pelos Estudos (Anos Finais)



(e) Nota – Infraestrutura da Escola (Anos Finais)



(f) Nota – Segurança da Escola (Anos Finais)

2.5.5. Programa Mais Educação e Indicadores Comportamentais

Uma hipótese que não podemos descartar é que a inclusão de atividades extracurriculares no contraturno - relacionadas a cultura, artes, esportes, lazer, entre outras - impacte positivamente a visão que os estudantes têm da escola, aumentando seu grau de satisfação. Ou ainda, que as atividades de acompanhamento pedagógico em português e matemática altere os hábitos dos alunos, criando hábitos de estudo e de leitura, ainda que essa melhora não se veja refletida em testes padronizados. Finalmente, o programa pode ter algum efeito no interesse dos alunos em continuar seus estudos, tendo por ambição ingressar em uma faculdade.

Buscamos avaliar essas hipóteses analisando o impacto do PME i) no percentual de alunos que alega gostar da escola; ii) no percentual de estudantes que indicaram ter o hábito da leitura em casa; iii) no percentual de estudantes que indicaram estudar em casa; iv) na nota que os alunos conferem à escola (de 0 a 10); e v) no percentual de alunos que indicam querer entrar em uma universidade no futuro. As duas últimas perguntas estão presentes somente no questionário dos Anos Finais.

Conforme fica evidente pela análise da Tabela 15, o programa não apresenta impacto significativo nem na satisfação com a escola nem em nenhum dos indicadores comportamentais avaliados. A ausência de impacto é ilustrada na Figura 10, que traz a evolução de alguns dos indicadores expostos na Tabela.

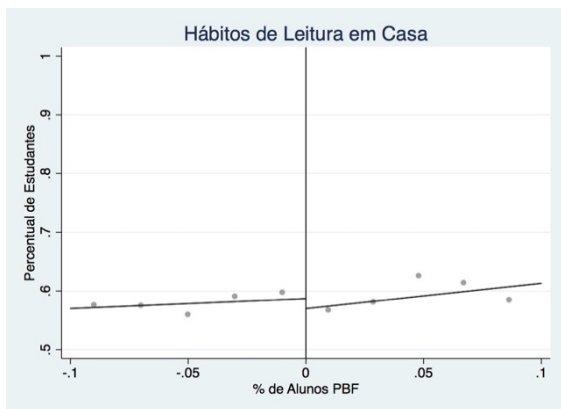
Tabela 15 – Efeitos do PME em indicadores comportamentais e na satisfação dos alunos com a escola (segundo estágio)

	Anos Iniciais					Anos Finais				
	Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i>		Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i>	
	5	10	15	Bandas Uniformes	Bandas Diferentes	5	10	15	Bandas Uniformes	Bandas Diferentes
Painel A: Gosta da Escola										
Coefficiente	-0.92	-0.20	-0.17	-0.21	-0.18	-0.13	0.00	0.08	-0.06	-0.04
p-valor	0.48	0.18	0.21	0.18	0.17	0.59	0.99	0.40	0.58	0.71
p-valor robusto				0.15	0.27				0.46	0.66
Intervalo				15	24 / 12				13	18 / 9
Painel B: Hábitos de Leitura em Casa										
Coefficiente	-0.66	-0.10	-0.04	-0.12	-0.09	0.24	0.17	0.12	0.14	0.13
p-valor	0.51	0.55	0.80	0.48	0.54	0.32	0.20	0.17	0.17	0.26
p-valor robusto				0.35	0.46				0.20	0.27
Intervalo				15	23 / 14				13	21 / 8
Painel C: Hábitos de Estudos em Casa										
Coefficiente	-1.01	-0.16	-0.10	-0.19	-0.18	0.10	0.11	0.11	0.08	0.09
p-valor	0.46	0.29	0.44	0.24	0.19	0.55	0.34	0.20	0.32	0.21
p-valor robusto				0.22	0.26				0.39	0.26
Intervalo				15	25 / 12				0.13	20 / 9
Painel D: Nota Escola										
Coefficiente						0.72	1.29	1.55	0.73	1.32
p-valor						0.69	0.29	0.13	0.40	0.16
p-valor robusto									0.58	0.21
Intervalo									14	19 / 8
Painel E: Interesse - Faculdade										
Coefficiente						-0.20	-0.14	0.05	-0.13	-0.12
p-valor						0.42	0.34	0.55	0.25	0.34
p-valor robusto									0.16	0.4
Intervalo									14	17 / 8

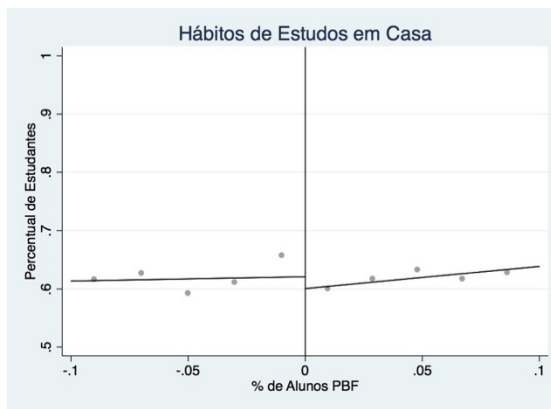
Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%.

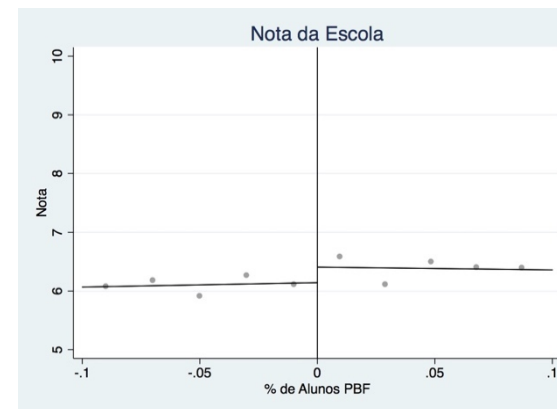
Figura 10 - Efeitos do PME em Indicadores Comportamentais e na Satisfação dos Estudantes com a Escola



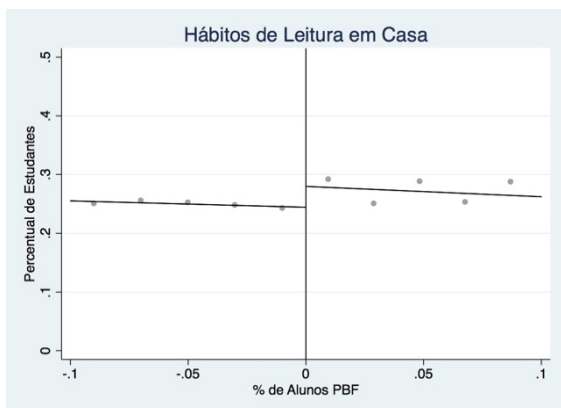
(a) Estudantes com hábitos de leitura (Anos Iniciais)



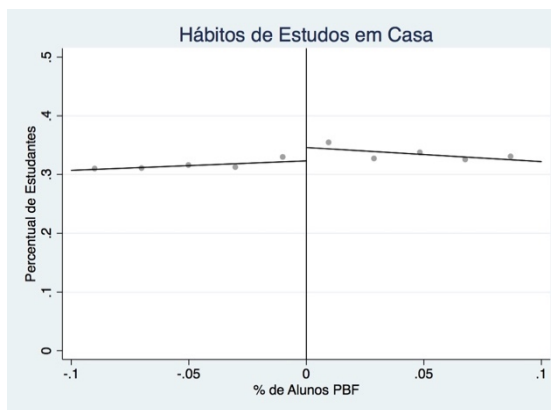
(b) Estudantes com hábitos de estudos (Anos Iniciais)



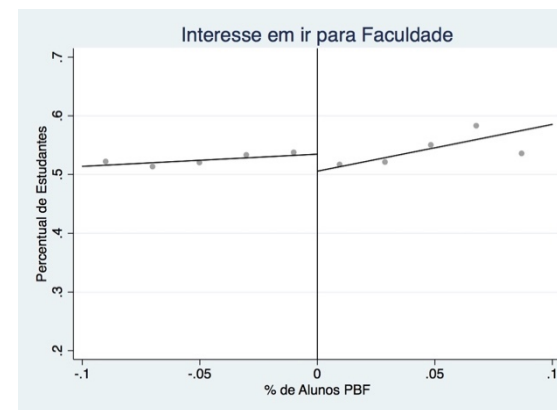
(c) Nota da Escola (Anos Finais)



(d) Estudantes com hábitos de leitura (Anos Finais)



(e) Estudantes com hábitos de estudos (Anos Finais)



(f) Interesse em fazer faculdade (Anos Finais)

2.5.6. Programa Mais Educação e Indicadores Acadêmicos

Conforme exposto na Introdução, diversos trabalhos se propuseram a avaliar o impacto do PME em indicadores educacionais⁵². Todos esses trabalhos utilizaram como referência o desempenho dos estudantes no Sistema de Avaliação da Educação Básica – Saeb. Nesta seção, iremos analisar o impacto do PME no desempenho acadêmico dos estudantes, medido por meio do percentual de estudantes que frequentaram aulas de recuperação em português, matemática, ciências e outras disciplinas durante o ano letivo. Trata-se de informação existente somente para estudantes dos Anos Finais.

Em coerência com estudos anteriores, não encontramos evidência de impacto significativo do PME nos indicadores acadêmico supracitados (vide Tabela 16 e Figura 11). Reforça-se, assim, a evidência de que o programa não impacta de forma significativa o desempenho acadêmico dos estudantes das escolas participantes nem mesmo em português e matemática, ainda que a única atividade obrigatória no âmbito do PME seja justamente o acompanhamento pedagógico nessas duas disciplinas.

⁵² Vide, por exemplo, Mendes (2011), Pereira (2011), Xerxenevsky (2012), Fundação Itaú (2015) e Oliveira & Terra (2016).

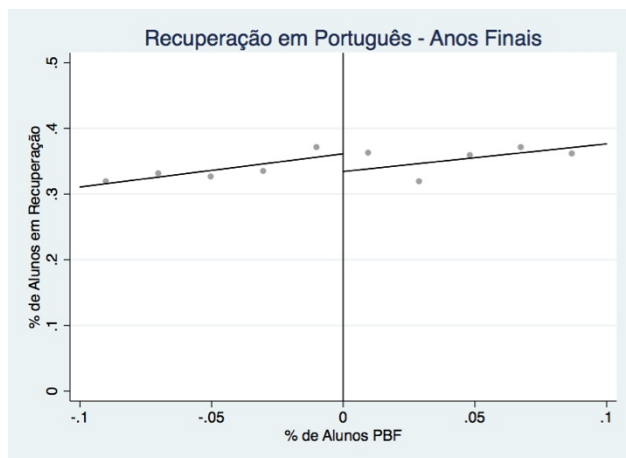
Tabela 16 – Efeitos do PME em indicadores acadêmicos (segundo estágio)

	Anos Finais				
	Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i>	
	5	10	15	Bandas Uniformes	Bandas Diferentes
Painel A: Recuperação - Português					
Coefficiente	-0.06	-0.13	0.01	-0.09	-0.02
p-valor	0.78	0.34	0.92	0.42	0.87
p-valor robusto				0.38	0.93
Intervalo				12	19 / 7
Painel B: Recuperação - Matemática					
Coefficiente	0.34	0.14	0.15	0.16	0.16
p-valor	0.39	0.43	0.31	0.27	0.34
p-valor robusto				0.32	0.33
Intervalo				14	18 / 7
Painel C: Recuperação Ciências					
Coefficiente	-0.11	-0.11	0.05	0.02	0.14
p-valor	0.70	0.48	0.64	0.87	0.32
p-valor robusto				1.00	0.24
Intervalo				11	19 / 7
Painel D: Recuperação Outros					
Coefficiente	-0.07	-0.05	0.03	0.06	0.15
p-valor	0.80	0.72	0.80	0.68	0.37
p-valor robusto				0.72	0.20
Intervalo				11	20 / 9

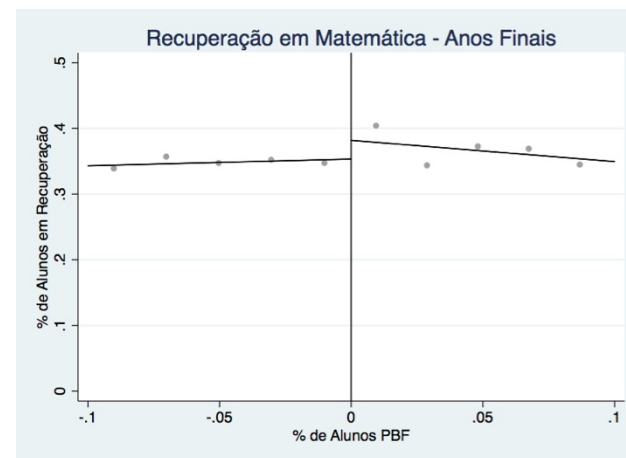
Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%.

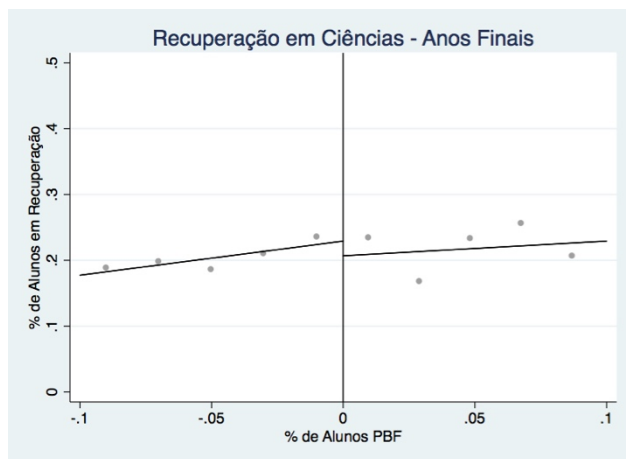
Figura 11 - Efeitos do PME em Indicadores Acadêmicos nos Anos Finais



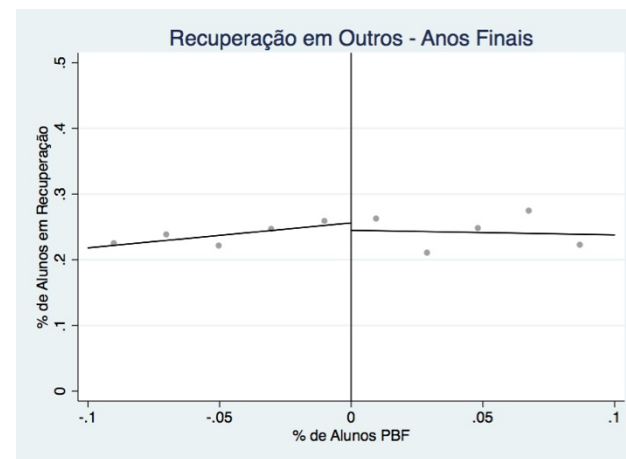
(a) Recuperação em Português (Anos Finais)



(b) Recuperação em Matemática (Anos Finais)



(c) Recuperação em Ciências (Anos Finais)



(d) Recuperação - Outros (Anos Finais)

2.6 Testes de Robustez

Nesta seção, realizaremos dois testes de falsificação/validação da estratégia empírica que são comumente aplicados em estimações por regressão descontínua. Esses testes visam conferir robustez aos resultados estimados.

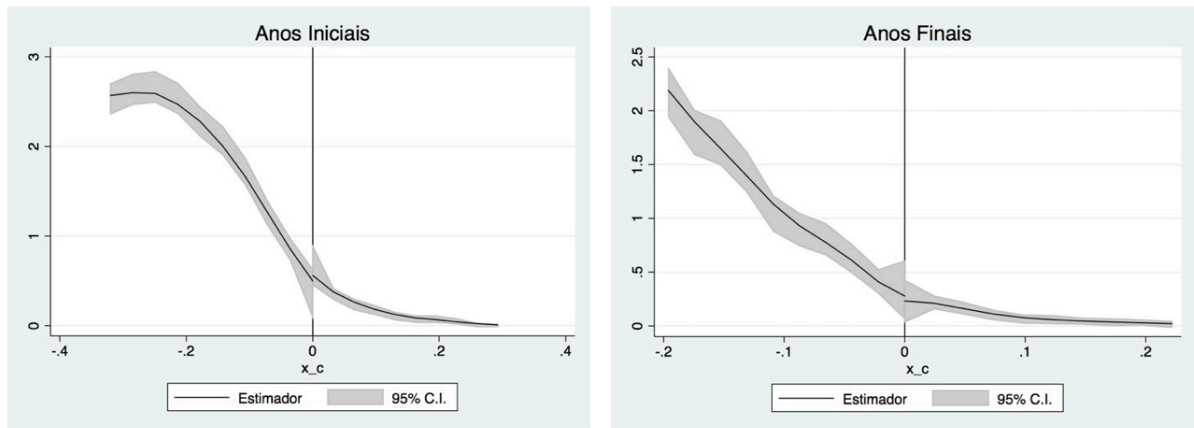
2.6.1. Teste de Densidade da Variável de Elegibilidade

Uma premissa fundamental em uma estratégia RDD é que características observáveis e não observáveis não variam de forma descontínua ao redor do *cutoff*. A designação para o tratamento ao redor do *cutoff* seria assim semelhante a um experimento aleatório (Lee, 2008). Isso não ocorreria caso as escolas pudessem, por exemplo, alterar o seu percentual de alunos PBF, de forma a se tornarem elegíveis para o tratamento. O Teste de McCrary visa justamente descartar a possibilidade de manipulação da variável de elegibilidade ao redor do *cutoff* (McCrary, 2008).

Intuitivamente, sabemos que é pouquíssimo provável que essa manipulação tenha ocorrido no caso do PME. Nossa variável de elegibilidade (percentual de alunos PBF na escola em 2010) somente foi adotada como critério de priorização do programa em 2011, tornando praticamente impossível sua manipulação. Assim, sabemos de antemão que a premissa de que os agentes não tenham controle sobre a variável de elegibilidade se sustenta.

Ainda assim, conduzimos teste formal proposto por Cattaneo, Jansson, & Ma (2017), em linha com o proposto por McCrary (2008). Os resultados podem ser observados na Figura 12. A ausência de manipulação é facilmente observada na análise gráfica, que representa a densidade do percentual de alunos PBF por escola. O painel (a) traz os testes para os Anos Iniciais e o painel (b) nos Anos Finais. Como se nota, em nenhum dos gráficos há sinal de descontinuidade na variável de elegibilidade, o que reforça a ausência de manipulação.

Figura 12 – Densidade da Variável de Elegibilidade



2.6.2. Testes Placebo e Diferença de Médias em Covariadas

A aplicação do teste de placebos em covariadas é outra forma de testar nossa premissa de “suavidade” da evolução das variáveis de interesse ao redor do *cutoff* na ausência do tratamento. Uma forma de reforçar essa premissa é testar para a existência de efeito do tratamento (descontinuidade) em variáveis que não deveriam ser por ele afetadas. Para isso, utilizaremos os mesmos métodos aplicados em variáveis de interesse em seções anteriores para avaliar uma série de covariadas que são, em tese, imunes ao tratamento. Essas variáveis podem ser agrupadas em três grandes grupos: perfil racial das famílias, escolaridade dos pais e infraestrutura das escolas. Testamos, no total, 15 covariadas⁵³.

Os resultados de algumas dessas estimações podem ser visualizados na Figura 13. Como se nota, não há evidência de descontinuidade significativa em nenhuma das covariadas. Os resultados completos desses testes de placebo podem ser conferidos na Tabela A.3 do Apêndice.

Adicionalmente, como vimos na seção 2.6.1, partimos do pressuposto que as escolas situadas ao redor do *cutoff* não diferem significativamente em termos de características observáveis e não-observáveis. Para avaliar essa premissa, conduzimos um teste de diferença de médias no grupo de escolas situados à esquerda e à direita do *cutoff*

⁵³ Não apresentamos os resultados para covariadas relacionadas à disponibilidade de água, energia e internet, usualmente testadas em trabalhos com o Censo Escolar, por tais serviços estarem praticamente universalizados em escolas urbanas do Estado de São Paulo.

para cada um dos intervalos arbitrariamente definidos - 5, 10 e 15. O objetivo é avaliar se as covariadas estão relativamente equilibradas nos dois grupos de escolas. As covariadas avaliadas foram as mesmas do teste de placebo. Os resultados podem ser observados na Tabela A.4 do Apêndice.

Como se nota, as covariadas apresentam diferenças significativas em toda a amostra, mas essas diferenças tendem a desaparecer à medida que diminuimos o intervalo de análise, o que reforça a ideia de uma distribuição aleatória do tratamento ao redor do *cutoff*. Nos dois intervalos mais curtos (5 e 10) praticamente não encontramos diferenças estatisticamente significativas entre as escolas antes e depois do *cutoff*, tanto para a base dos Anos Iniciais quanto para a base dos Anos Finais. Já o intervalo de 15 acaba por abarcar escolas distintas entre si em alguns aspectos, tais como o número de salas existentes e a escolaridade média dos pais dos alunos. Por essa razão, entendemos que o intervalo de 10 é o que melhor se adequa ao nosso objetivo de nos manter tão perto quanto possível do *cutoff*, ao mesmo tempo em que permanecemos com um número razoável de observações para estimação e inferência.

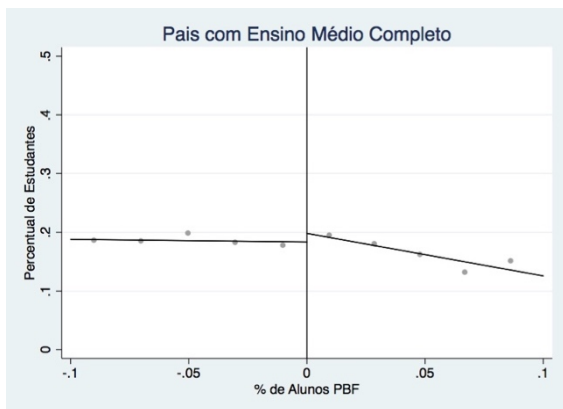
Os testes conduzidos corroboram, assim, a validade da estratégia empírica adotada.

2.7 Estratégia Empírica Alternativa: Uma Estimação por Diferenças-em-Diferenças

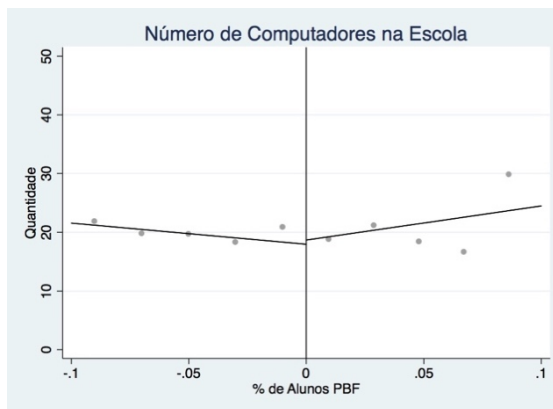
2.7.1. Metodologia

Uma limitação relevante da estimação por regressão descontínua diz respeito à validade externa. Isto é, o impacto estimado não é generalizável à todas escolas, mas restrito às escolas situadas ao redor do *cutoff*. Essa limitação ocorre em detrimento de um grande ganho em termos de validade interna. Isto porque a metodologia é bastante robusta no sentido de estimar a causalidade do programa nas escolas situadas nos intervalos definidos ao redor do *cutoff*.

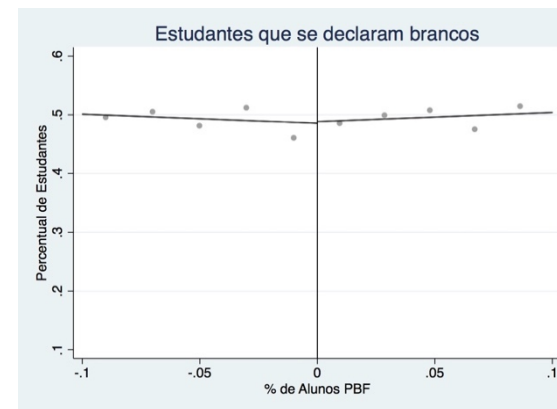
Figura 13 - Efeitos Placebos do PME (testes em covariadas)



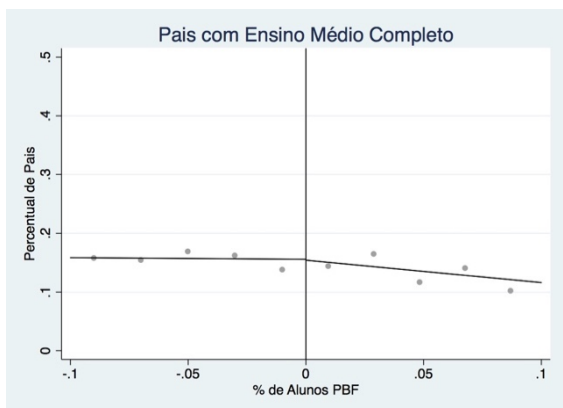
(a) Pais com Ensino Médio Completo – Anos Iniciais



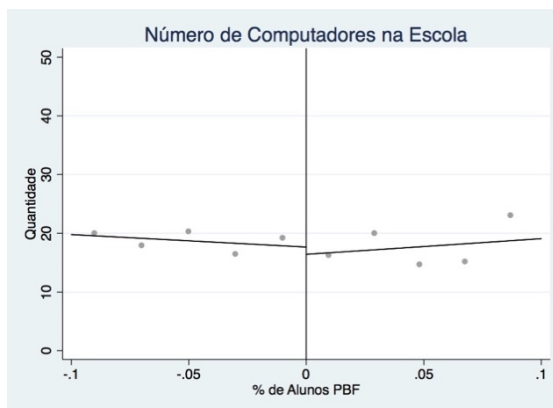
(b) Número de Computadores na Escola - Anos Iniciais



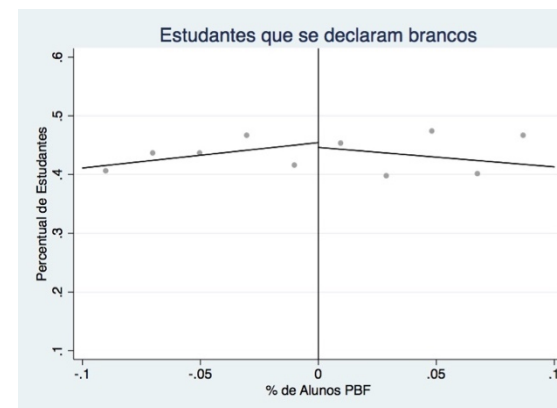
(c) Estudantes que se declaram brancos - Anos Iniciais



(d) Pais com Ensino Médio Completo – Anos Finais



(e) Número de Computadores na Escola - Anos Finais



(f) Estudantes que se declaram brancos - Anos Finais

Como forma de contornar essa limitação, tornando os resultados do artigo mais generalizável, complementaremos a estimação realizada nas seções anteriores com uma estratégia empírica baseada no método de diferenças-em-diferenças. De forma bastante simples, este método consiste em comparar amostras de escolas tratadas e não tratadas antes e depois da intervenção. A estratégia consiste em tirar “duas diferenças”: a diferenças nas médias da variável de interesse para os grupos de tratamento e controle antes do tratamento e a mesma diferença depois do tratamento. Dessa forma, esperamos controlar para a presença de fatores externos que podem afetar os dois grupos no período⁵⁴.

Nota-se que, pelo método, construímos o contrafactual tendo por premissa o fato de que o viés de seleção não varia com o tempo. Trata-se de uma importante limitação dessa metodologia. Logo, vamos partir do pressuposto de que as mudanças na variável de interesse do grupo controle revele o que teria acontecido com o grupo tratado na ausência do tratamento:

$$E(Y_1^T - Y_0^T | T_1 = 0) = E(Y_1^C - Y_0^C | T_1 = 0) \quad (4)$$

onde Y^T designa a variável de interesse para o grupo tratado, Y^C para o grupo controle e T o *status* do tratamento.

Como no período 0 nenhum grupo foi tratado, temos que $T_{0i}=0$ e $Y_{0i}=Y_{0i}^C$ para todo i . O estimador de DD, que mede o impacto do tratamento nos tratados no período 1, é dado por:

$$DD = E(Y_1^T - Y_0^T | T_1 = 1) - E(Y_1^C - Y_0^C | T_1 = 0) = E(I_1 | T_1 = 1) \quad (5)$$

O primeiro elemento do lado esquerdo da equação traz a diferença na variável de interesse para o grupo dos tratados antes (Y_0^T) e depois (Y_1^T) do tratamento; já o segundo elemento, a diferença na variável de interesse para o grupo dos controles antes (Y_0^C) e depois (Y_1^C) do tratamento. O estimador de DD é, portanto, a diferença entre esses dois elementos.

Como vimos, a técnica de diferença-em-diferença não controle para fatores que afetem de forma desigual os grupos de tratamento e controle. Para diminuir a

⁵⁴ Para a apresentação formal do método e sua aplicação, vide Card & Krueger (1994).

probabilidade de que isso ocorra, vamos combinar essa metodologia com a técnica de pareamento por escore de propensão - PSM. Essa técnica consiste em selecionar um conjunto de características observáveis e, a partir dessas características, encontrar para cada escola que participa do PME uma escola que não participa com características similares⁵⁵.

Conforme destacam Hirano, Imbens, & Ridder (2003), o uso conjunto das duas técnicas possibilita a estimação de uma regressão ponderada pelo escore de propensão de cada unidade, corrigindo distorções do modelo DD. Dessa forma, controlamos para características observáveis com o PSM e para variáveis não observáveis invariantes no tempo utilizando a DD⁵⁶.

A combinação das duas técnicas foi realizada em dois passos. Em primeiro lugar, realizamos o pareamento com base em características das escolas antes do tratamento – utilizamos como base o ano de 2011. Utilizamos para o pareamento cerca de 40 variáveis divididas em quatro grandes grupos: características acadêmicas da escola, características físicas da escola, características socioeconômicas das famílias e características do município onde a escola está situada. Os resultados do pareamento pelo vizinho mais próximo podem ser consultados na Tabela A.5⁵⁷.

Como se nota, antes do pareamento, os grupos tratamento apresentavam diferenças estatisticamente significativas na grande maioria das características, o que indica as vantagens de se proceder o pareamento (do contrário, estaríamos comparando escolas bastante diferentes entre si). Após o pareamento, tanto para os Anos Iniciais quanto para os Anos Finais, não há nenhuma diferença significativa entre as escolas do grupo tratamento e as escolas do grupo controle pareado. Logo, o pareamento foi realizado com êxito.

⁵⁵ Para a apresentação formal do método e sua aplicação, vide Rosenbaum & Rubin (1983).

⁵⁶ Note que a combinação ainda possui limitações, uma vez que não é controla para características não observadas que afetem desigualmente o grupo de tratamento e de controle. Contudo, o pareamento diminui a probabilidade de que isso ocorra, uma vez que cria um grupo controle com características bastante similares a do grupo tratamento.

⁵⁷ Existem várias técnicas para realização do pareamento: pareamento pelo vizinho mais próximo (*nearest-neighbor matching*), pelos 5 vizinhos mais próximos (*nearest 5-neighbor matching*), pareamento por raio (*radius matching*) e pareamento kernel (*kernel matching*). Neste trabalho realizamos as estimações pelos dois métodos mais comuns na literatura: vizinho mais próximo com reposição e *kernel*.

Após o pareamento, aplicamos o método de diferenças-em-diferenças somente nas unidades pareadas.

2.7.2. Impacto nas Variáveis de Interesse

A Tabela 17 traz os resultados das estimativas do impacto do PME no emprego das mães e na satisfação dos pais com a escola. A Tabela 18, os resultados do impacto do programa para os demais indicadores avaliados.

Os coeficientes estimados consistem nas diferenças na evolução das variáveis de interesse dos grupos tratamento e controle entre 2011 e 2012 – ou seja, são o estimador DD. A tabela reporta o impacto do PME utilizando três métodos: diferenças-em-diferenças (DD) sem pareamento, diferenças-em-diferenças com pareamento *kernel* (PSM + DD – *kernel*) e com pareamento pelo método do vizinho mais próximo – *nearest neighbour* (PSM + DD – NN).

Como se nota, assim como na estimação por RDD, não se observa qualquer impacto significativo do Programa em nenhuma das variáveis avaliadas. Em alguns casos, verifica-se impacto estatisticamente significativo na comparação entre o grupo tratamento e controle sem pareamento. Contudo, os resultados não se mostram significativos nas estimações mais robustas, realizadas com pareamento.

Logo, os resultados estimados no artigo mostram-se robustos a escolha do método de estimação, o que reforça a ausência de impacto significativo do Programa Mais Educação em indicadores sociais, comportamentais e na satisfação de pais e alunos com a escola.

Tabela 17– Efeitos do PME no Emprego das Mães e na Satisfação dos Pais com a Escola (evolução dos indicadores entre 2011 e 2012)

Variável	Anos Iniciais			Anos Finais		
	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)
Impacto no Emprego/Desemprego da Mãe						
△ Mãe Empregada	0.007 (0.006)	0.000 (0.006)	0.003 (0.009)	0.006 (0.006)	0.002 (0.006)	0.000 (0.008)
△ Mãe Desempregada	0.011** (0.005)	0.014* (0.006)	0.013 (0.009)	0.004 (0.004)	0.000 (0.005)	0.003 (0.006)
△ Mãe Emprego Temporário	-0.001 (0.003)	-0.002 (0.003)	0.004 (0.004)	0.000 (0.002)	-0.001 (0.002)	-0.001 (0.003)
△ Mãe Autônoma	0.004 (0.004)	0.001 (0.005)	0.003 (0.007)	0.004 (0.004)	0.003 (0.004)	0.001 (0.006)
△ Mãe Negócio Próprio	-0.001 (0.002)	-0.002 (0.002)	0.001 (0.003)	0.001 (0.002)	0.001 (0.002)	0.001 (0.003)
Impacto na Satisfação dos Pais com a Escola						
△ Nota - Professores da Escola	0.009 (0.026)	0.002 (0.028)	-0.019 (0.04)	-0.002 (0.027)	-0.001 (0.027)	-0.055 (0.042)
△ Nota - Infraestrutura da Escola	0.071 (0.047)	0.055 (0.064)	-0.014 (0.089)	0.016 (0.048)	0.007 (0.054)	0.050 (0.068)
△ Nota - Segurança da Escola	0.022 (0.051)	0.044 (0.062)	-0.014 (0.083)	0.103* (0.055)	0.073 (0.062)	0.042 (0.082)
△ Nota - Interesse nos Estudos	-0.010 (0.03)	-0.008 (0.033)	-0.034 (0.049)	0.005 (0.027)	-0.002 (0.028)	-0.003 (0.037)

Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%

Tabela 18– Efeitos do PME em Indicadores Comportamentais, na Satisfação dos Estudantes com a Escola, no Trabalho Infantil e em Indicadores Acadêmicos

Variável	Anos Iniciais			Anos Finais		
	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)
Impacto em Indicadores Comportamentais e na Satisfação dos Estudantes com a Escola						
△ Estudantes que gostam da escola	0.015* (0.008)	0.012 (0.008)	0.004 (0.013)	0.006 (0.007)	0.001 (0.007)	-0.003 (0.009)
△ Estudantes com Hábitos de Leitura	0.011 (0.009)	0.015 (0.01)	0.013 (0.014)	0.004 (0.005)	0.007 (0.005)	0.006 (0.007)
△ Estudantes com Hábitos de Estudos	0.010 (0.009)	0.010 (0.01)	0.013 (0.014)	0.001 (0.006)	0.001 (0.006)	-0.007 (0.008)
△ Nota da Escola				0.11** (0.016)	0.056 (0.051)	0.019 (0.073)
△ Interesse em ir para a Faculdade				-0.009 (0.006)	-0.010 (0.006)	-0.004 (0.008)
Impacto no Trabalho Infantil						
△ Trabalho Doméstico				0.007 (0.006)	0.006 (0.006)	0.008 (0.009)
△ Trabalho fora de casa				-0.002 (0.003)	-0.004 (0.003)	-0.004 (0.005)
Impacto em Indicadores Acadêmicos						
△ Recuperação em Português				0.014** (0.001)	0.012 (0.008)	0.003 (0.011)
△ Recuperação em Matemática				0.018** (0.11)	0.012 (0.009)	0.000 (0.013)

Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%

2.8 Conclusão

A implementação de programas de educação integral entrou, em anos recentes, na pauta dos mais diferentes níveis de governo e de organizações da sociedade civil. Um sintoma da importância desse tema para a sociedade brasileira é a sua onipresença na plataforma política de candidatos em períodos eleitorais. A título de exemplo, todos os candidatos a presidente em 2014 afirmaram defender a ampliação da “educação integral” no país⁵⁸. Esse consenso entre candidatos reflete, certamente, a percepção de que a grande maioria da população apoia a ampliação de programas voltados ao aumento da jornada escolar no país.

Não é raro, portanto, que a educação integral seja apresentada como verdadeira panaceia para os problemas sociais e educacionais no País. Nesse contexto, julgamos de suma importância que a academia desempenhe seu papel no sentido de informar à sociedade o real impacto de certas políticas públicas *vis a vis* seu custo.

Com relação ao desempenho acadêmico, diversos trabalhos vêm consolidando o entendimento acerca da ausência de impacto do PME. Contudo, o discurso a favor da educação integral muitas vezes aborda não somente seu aspecto acadêmico, mas também seu aspecto social, na medida em que, supostamente, facilitaria a inserção das mães no mercado de trabalho, ao mesmo tempo em que diminuiria o trabalho infantil. Neste trabalho, buscamos complementar as avaliações já existentes do Programa Mais Educação, focadas no impacto do programa no desempenho acadêmico dos estudantes, realizando a primeira avaliação de impacto do PME em indicadores sociais (desemprego da mãe e trabalho infantil) e de satisfação de pais e alunos com a escola.

Em síntese, vimos que o programa não impactou de forma significativa nenhuma das variáveis avaliadas. Chama a atenção a ausência de impacto do programa, particularmente no desemprego da mãe, no trabalho infantil e na satisfação dos alunos com a escola. Como vimos na seção 2.2, um dos critérios mais importantes de priorização do programa é a vulnerabilidade social das famílias dos estudantes da escola. Assim, seria

⁵⁸ Vide, por exemplo, as seguintes referências aos programas dos candidatos: i) <http://portal.aprendiz.uol.com.br/2014/09/25/o-que-pensam-os-presidenciais-sobre-educacao-integral/>; ii) <http://noticias.r7.com/eleicoes-2014/educacao-integral-conheca-as-propostas-dos-candidatos-a-presidencia-02102014>; iii) <http://www.ebc.com.br/noticias/eleicoes-2014/2014/09/educacao-integral-conheca-as-propostas-dos-candidatos-a-presidencia>. Acessados em 14 de julho de 2017.

de se esperar que, ao praticamente dobrar o tempo de permanência do estudante na escola, o programa impactasse positivamente a empregabilidade da mãe, bem como impactasse negativamente o trabalho infantil, diminuindo o percentual dos jovens que estudam e trabalham. Adicionalmente, ao aumentar o leque de atividades desenvolvidas pela escola (incluindo, por exemplo, atividades de cultura e artes, esportes e lazer, cultura digital, etc.), esperaríamos que o PME aumentasse a satisfação dos pais e dos estudantes com a escola, bem como mudasse alguns hábitos, estimulando o estudo e a leitura. Contudo, em nenhum desses casos vimos um impacto significativo do programa.

É importante ressaltar que os resultados deste trabalho, bem como de outros que demonstram a ausência de impacto do PME no desempenho acadêmico, não representam um argumento contra políticas de educação integral *per se*. Não temos dúvidas de que a expansão da jornada escolar *pode* contribuir para o aumento da qualidade da educação no País, bem como para a melhora de certos indicadores sociais. O que esses estudos mostram é que a educação integral, *da forma como vem sendo estimulada e financiada pelo Governo Federal no âmbito do Programa Mais Educação*, não vem apresentando resultados satisfatórios para a sociedade brasileira.

Finalmente, julgamos fundamental que análises de programas como o Mais Educação não fiquem limitadas ao seu impacto no desempenho acadêmico dos estudantes. É indiscutível que a escola constitui parte fundamental da vida de crianças, jovens e de suas famílias, exercendo uma influência que vai muito além do desenvolvimento acadêmico. Esperamos que este trabalho estimule outros pesquisadores a avaliar o impacto da educação integral e de outros programas educacionais em uma série de outros indicadores, tais como criminalidade, comportamentos de risco, entre outros.

2.9 Referências Bibliográficas

- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics : an empiricist's companion*. Princeton University Press.
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, *82*(6), 2295–2326. <http://doi.org/10.3982/ECTA11757>
- Card, D., & Krueger, A. (1994). Minimum Wages and Employment: A Case Study of the Fast- Food Industry in New Jersey and Pennsylvania. *The American Economic Review*, *84*, 772–793.
- Cattaneo, M. D., Jansson, M., & Ma, X. (2017). *Simple Local Polynomial Density Estimators*. Retrieved from <https://sites.google.com/site/rdpackages/rddensity>
- Cattaneo, M. D., & Titiunik, R. (2017). *A Practical Introduction to Regression Discontinuity Designs*. Retrieved from http://www-personal.umich.edu/~cattaneo/books/Cattaneo-Idrobo-Titiunik_2017_Cambridge.pdf
- Cattaneo, M. D., & Vazquez-Bare, G. (2016). The Choice of Neighborhood in Regression Discontinuity Designs. *Observational Studies*, *2*, 134–146. Retrieved from http://www-personal.umich.edu/~cattaneo/papers/Cattaneo-VazquezBare_2016_ObsStud.pdf
- Fundação Itaú. (2015). O Programa Mais Educação. Relatório de Avaliação Econômica e Estudos Qualitativos. *Fundação Itaú Social*. Retrieved from https://www.redeitausocialdeavaliacao.org.br/wp-content/uploads/2015/11/relat_Mais_Educacao_COMPLETO_20151118.pdf
- Gelman, A., & Imbens, G. (2014). *Why High-order Polynomials Should not be Used in Regression Discontinuity Designs*. Cambridge, MA. <http://doi.org/10.3386/w20405>
- Hahn, J., Todd, P., & Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209. <http://doi.org/10.1111/1468-0262.00183>
- Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica*, *71*(4), 1161–1189. <http://doi.org/10.1111/1468-0262.00442>
- Imbens, G., & Kalyanaraman, K. (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies*, *79*(3), 933–959. <http://doi.org/10.1093/restud/rdr043>

- Jacob, B., & Lefgren, L. (2003). *Are Idle Hands the Devil's Workshop? Incapacitation, Concentration and Juvenile Crime*. Cambridge, MA. <http://doi.org/10.3386/w9653>
- Kruger, D. I., & Berthelon, M. E. (2009). Delaying the Bell: The Effects of Longer School Days on Adolescent Motherhood in Chile. *Discussion Paper No. 4553*, (4553).
- Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, *142*(2), 675–697. <http://doi.org/10.1016/j.jeconom.2007.05.004>
- Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281–355. <http://doi.org/10.1257/jel.48.2.281>
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714. <http://doi.org/10.1016/j.jeconom.2007.05.005>
- MEC. (2014). *Manual Operacional da Educação Integral*.
- Melo, D. M. (2015). Consultorias de 2013 a 2015 - Resumo Executivo. *Projeto 914BRZ2003/2013. Projeto 914BRZ2002/2014*.
- Mendes, K. M. (2011). *O Impacto do Programa Mais Educação no Desempenho dos Alunos da Rede Pública Brasileira*. Universidade de São Paulo.
- Oliveira, L. F. B. de, & Terra, R. (2016). Impacto do Programa Mais Educação em indicadores educacionais. *Working Paper - International Policy Centre for Inclusive Growth*. Retrieved from http://www.ipc-undp.org/pub/port/WP147PT_Impacto_do_Programa_Mais_Educacao.pdf
- Pereira, G. C. (2011). *Uma avaliação de impacto do programa Mais Educação no Ensino Fundamental*. Universidade Federal do Rio de Janeiro.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*(1), 41. <http://doi.org/10.2307/2335942>
- Xerxenevsky, L. L. (2012). *Programa Mais Educação: avaliação do impacto da educação integral no desempenho de alunos no Rio Grande do Sul*. Pontifícia Universidade Católica do Rio Grande do Sul.

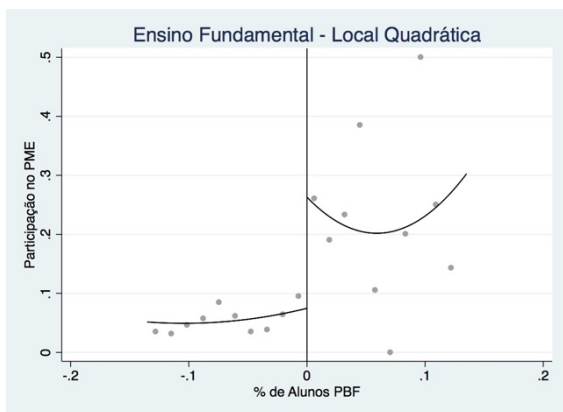
2.10 Apêndice

Tabela A.1. – Macrocâmpos e Exemplos de Atividades do Programa Mais Educação em Escolas Urbanas⁵⁹

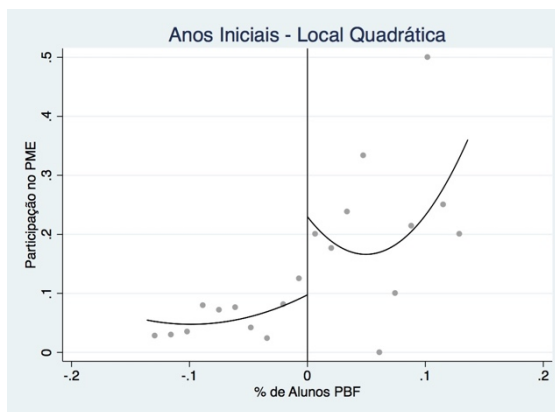
Macrocampo	Atividades
Acompanhamento Pedagógico (obrigatório)	<ul style="list-style-type: none"> • Orientação de Estudos e Leituras (que contempla diferentes áreas do conhecimento: alfabetização, matemática, história, línguas estrangeiras, entre outras); • Deve ser realizada com duração de uma hora a uma hora e meia diariamente.
Comunicação, Uso de Mídias e Cultura Digital e Tecnológica	<ul style="list-style-type: none"> • Promoção da Saúde; • Educação em Direitos Humanos; • Ambiente de Redes Sociais; • Fotografia; • História em Quadrinhos; • Jornal e Rádio Escolares; • Robótica Educacional; • Entre outros.
Cultura, Artes e Educação Patrimonial	<ul style="list-style-type: none"> • Canto Coral; • Capoeira; • Danças; • Pintura; • Teatro; • Entre outros.
Educação Ambiental, Desenvolvimento Sustentável e Economia Solidária e Criativa/Educação Econômica	<ul style="list-style-type: none"> • Horta Escolar e/ou Comunitária; • Jardinagem Escolar; • Economia Solidária e Criativa/Educação Econômica;
Esporte e Lazer	<ul style="list-style-type: none"> • Atletismo; • Basquete; • Futebol; • Ginástica Rítmica; • Judô; • Yoga/Meditação; • Entre outros.
Educação em Direitos Humanos	<ul style="list-style-type: none"> • Educação em Direitos Humanos.
Promoção da Saúde	<ul style="list-style-type: none"> • Promoção da Saúde e Prevenção de Doenças e Agravos à Saúde.

⁵⁹ Para mais detalhes sobre os macrocampos e suas atividades, vide (MEC, 2014). Para os macrocampos e atividades de escolas do campo, vide página 21 e 22 do referido documento ou o §6º do art. 3º da Resolução nº 14, de 9 de junho de 2014, do Conselho Deliberativo do Fundo Nacional de Desenvolvimento da Educação.

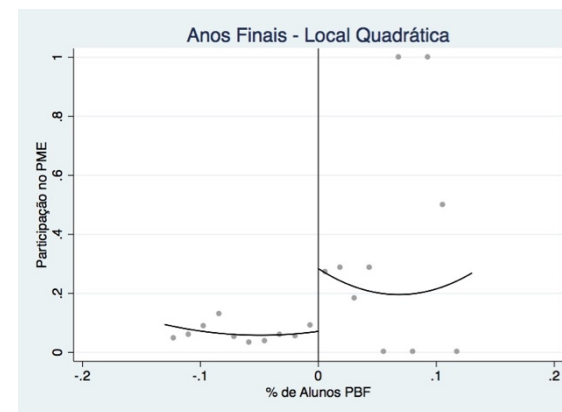
Figura A.1 - Descontinuidade na Participação no PME em 2012 (aproximação quadrática)



(a) Ensino Fundamental (local quadrática)



(b) Anos Iniciais (local quadrática)



(c) Anos Finais (local quadrática)

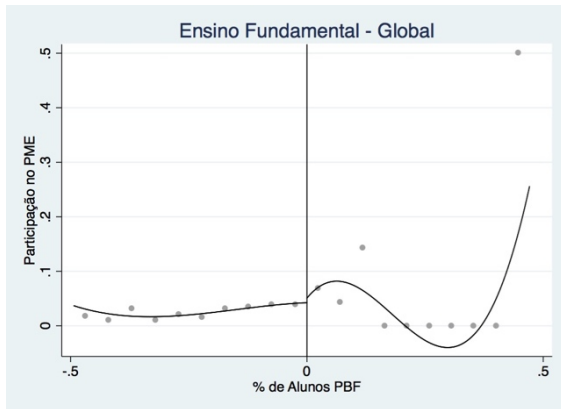
Tabela A.2 – Estimativas dos Efeitos da Elegibilidade para o PME na Participação no Programa em 2011 (primeiro estágio)

	Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i> ($p=1$)		Intervalos <i>Data-Driven</i> ($p=2$)	
	5	10	15	Bandas Uniformes	Bandas Diferentes	Bandas Uniformes	Bandas Diferentes
Média dos							
Controles							
Painel A: Ensino Fundamental							
Coefficiente	0.01	0.01	0.00	0.00	0.00	0.01	0.00
p-valor	0.78	0.89	0.89	0.93	0.93	0.89	0.94
p-valor robusto	-	-	-	0.90	0.89	0.78	0.84
Intervalo	-	-	-	11	20 / 11	16	27 / 16
Observações à esquerda/direita	167 / 113	495 / 168	1011 / 198	593 / 178	1662 / 178	1116 / 199	2789 / 199
Painel B: Anos Iniciais							
Coefficiente	0.01	0.03	0.03	0.02	0.02	0.04	0.03
p-valor	0.84	0.41	0.41	0.63	0.59	0.51	0.51
p-valor robusto	-	-	-	0.76	0.65	0.48	0.44
Intervalo	-	-	-	14	17 / 14	15	24 / 15
Observações à esquerda/direita	133 / 89	378 / 134	749 / 158	671 / 153	902 / 151	712 / 155	1583 / 157
Painel C: Anos Finais							
Coefficiente	-0.02	-0.04	-0.03	0.02	-0.01	0.03	-0.03
p-valor	0.83	0.53	0.60	0.84	0.81	0.79	0.70
p-valor robusto	-	-	-	0.73	0.76	0.64	0.61
Intervalo	-	-	-	5	25 / 6	9	21 / 8
Observações à esquerda/direita	72 / 37	216 / 53	443 / 64	72 / 37	1242 / 44	181 / 52	812 / 51

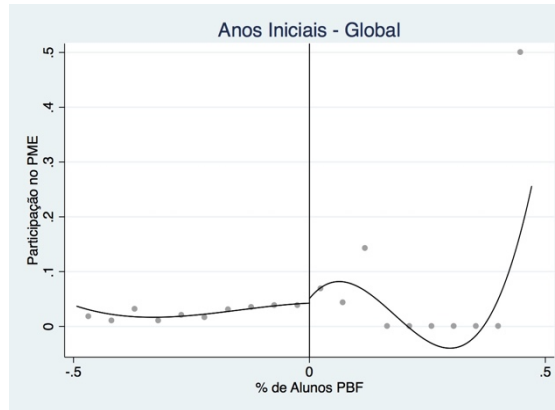
Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%. * Significativo a 10%

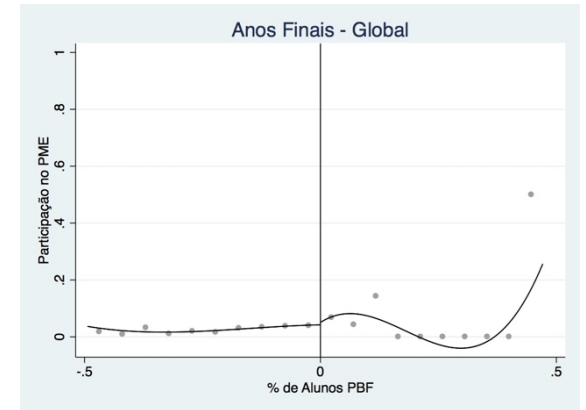
Figura A.2 - Descontinuidade na Participação no PME em 2011



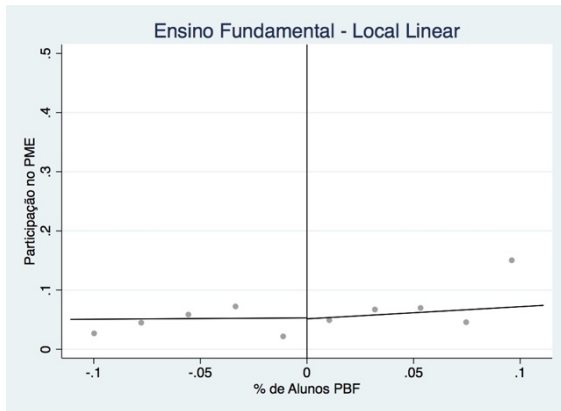
(a) Ensino Fundamental - Global



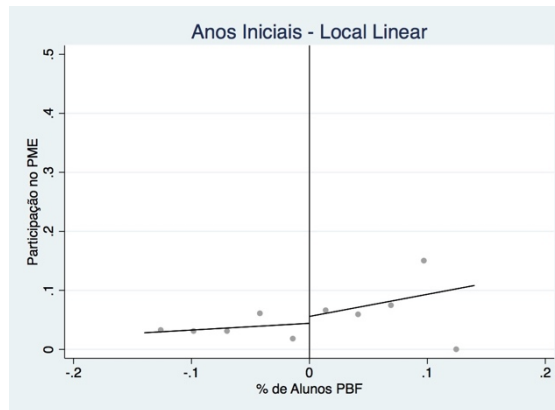
(c) Anos Iniciais - Global



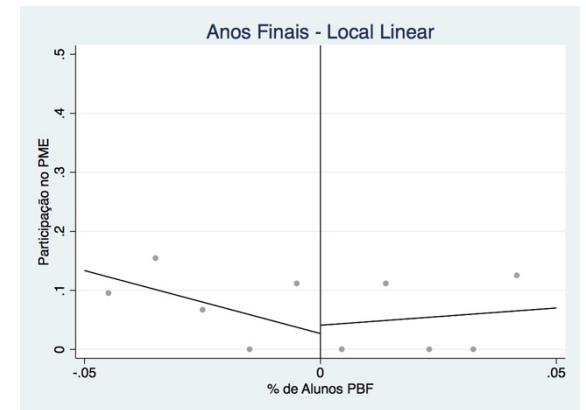
(d) Anos Finais - Global



(d) Ensino Fundamental - Linear Local



(e) Anos Iniciais - Linear Local



(f) Anos Finais - Linear Local

Tabela A.3 – Efeitos Placebo do PME (testes em covariadas) (primeiro estágio)

	Intercepto dos Controles	Anos Iniciais				Intercepto dos Controles	Anos Finais			
		Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i> Bandas Uniformes		Intervalos <i>Ad Hoc</i>			Intervalos <i>Data-Driven</i> Bandas Uniformes
		5	10	15			5	10	15	
Painel A: Alunos Brancos										
Coefficiente	0.49	0.02	0.00	0.02	0.00	0.45	0.06	0.00	0.01	0.07
p-valor		0.49	0.91	0.39	1.00		0.15	0.91	0.81	0.13
p-valor robusto					0.98					0.10
Intervalo					0.10					0.03
Painel B: Alunos Negros										
Coefficiente	0.10	-0.03***	-	-	-0.01	0.10	-0.03	-0.01	0.00	-0.03
p-valor		0.01	0.21	0.10	0.34		0.13	0.74	0.73	0.13
p-valor robusto					0.37					0.10
Intervalo					0.07					0.04
Painel C: Pai - Ensino Fundamental										
Coefficiente	0.19	-0.01	0.01	0.01	0.01	0.24	-0.02	-0.03	-0.02	0.00
p-valor		0.51	0.19	0.18	0.63		0.38	0.14	0.17	1.00
p-valor robusto					0.86					0.84
Intervalo					0.11					0.04
Painel D: Pai - Ensino Médio										
Coefficiente	0.18	0.04	0.01	0.02	0.02	0.16	0.02	0.00	0.01	0.01
p-valor		0.08	0.29	0.07	0.16		0.34	0.95	0.56	0.40
p-valor robusto					0.24					0.38
Intervalo					0.11					0.06
Painel E: Pai - Ensino Superior										
Coefficiente	0.02	0.01	0.01	0.01	0.01	0.03	0.00	0.00	0.00	0.00

p-valor		0.10	0.13	0.15			0.07		0.84	0.73	0.84		0.82
p-valor robusto							0.07						0.79
Intervalo							0.08						0.06

Painel F: N° de Alunos (Iniciais/Finais)

			-	-				-32.6	-80.5	-20			
Coefficiente	315.09	-35.04	15.82	5.99	28.65	319.49		0.63	0.13	0.63			-18.84
p-valor		0.54	0.70	0.84	0.28			0.92	0.83	0.92			0.48
p-valor robusto					0.31								0.45
Intervalo					0.10								0.08

Painel G: Número de Funcionários

			-	-				0.99	-1.65	0.58			
Coefficiente	44.22	-2.26	3.30	2.04	2.80	52.16		0.92	0.83	0.92			2.77
p-valor		0.66	0.36	0.48	0.19			0.98	0.06	0.32			0.59
p-valor robusto					0.16								0.69
Intervalo					0.12								0.07

Painel H: Número de Salas Existentes

			-	-				0.03	-2.2	0.92			
Coefficiente	10.96	-0.52	1.12	0.74	0.23	11.09		0.98	0.06	0.32			0.84
p-valor		0.72	0.20	0.29	0.76			0.98	0.06	0.32			0.45
p-valor robusto					0.85								0.34
Intervalo					0.09								0.05

Painel I: Número de Computadores

								-2.31	-1.2	-1.6			-1.69
Coefficiente	17.97	-2.56	0.71	0.49	-0.97	17.67		0.54	0.60	0.32			0.58
p-valor		0.55	0.81	0.82	0.77			0.54	0.60	0.32			0.58
p-valor robusto					0.64								0.50
Intervalo					0.08								0.07

Painel J: Biblioteca

								0.03	0.03	0.02			
Coefficiente	0.06	0.28***	0.14	0.05	0.13	0.09		0.03	0.03	0.02			0.06
p-valor		0.01	0.06	0.42	0.07			0.81	0.79	0.85			0.49
p-valor robusto					0.07								0.60

Intervalo					0.10					0.06
Painel K: Refeitório										
Coefficiente	0.46	0.21	0.08	0.06	0.16	0.49	0.30	0.19	0.17	0.28
p-valor		0.13	0.43	0.48	0.13		0.13	0.20	0.16	0.09
p-valor robusto					0.12					0.09
Intervalo					0.09					0.10
Painel L: Sala Professores										
Coefficiente	0.90	0.03	0.04	0.02	0.03	0.93	0.07	0.13**	0.02	0.07
p-valor		0.64	0.39	0.57	0.43		0.30	0.03	0.55	0.15
p-valor robusto					0.37					0.32
Intervalo					0.15					0.05
Painel M: Laboratório de Informática										
Coefficiente	0.88	-0.07	0.01	0.01	0.00	0.86	0.12	0.16**	0.06	0.08
p-valor		0.37	0.87	0.83	0.97		0.19	0.02	0.17	0.28
p-valor robusto					1.00					0.44
Intervalo					0.11					0.06
Painel N: Laboratório de Ciências										
Coefficiente	0.03	-0.05	0.01	0.04	-0.01	0.08	0.02	0.04	0.11	0.08
p-valor		0.40	0.82	0.08	0.78		0.88	0.71	0.18	0.55
p-valor robusto					0.53					0.65
Intervalo					0.11					0.08
Painel O: Quadra de Esportes										
Coefficiente	0.72	0.14	0.05	0.01	0.17	0.89	-0.07	0.08	0.12**	0.02
p-valor		0.23	0.54	0.83	0.07		0.26	0.16	0.04	0.64
p-valor robusto					0.07					0.87
Intervalo					0.08					0.06

*** Significativo a 1% ** Significativo a 5%

Tabela A.4 – Diferenças de médias entre escolas minoria PBF e escolas maioria PBF em diferentes intervalos

	Média dos Controles (intervalo de 10)	Anos Iniciais				Média dos Controles (intervalo de 10)	Anos Finais			
		Intervalos <i>Ad Hoc</i>					Intervalos <i>Ad Hoc</i>			
		5	10	15	Global		5	10	15	Global
Observações à esquerda/direita		128/83	365/126	725 / 147	3569 / 176		66 / 36	206 / 52	422 / 61	
Painel A: Alunos Brancos										
Coefficiente	0.49	0.00	0.00	-0.02	-0.05***	0.43	-0.02	0.01	0.01	-0.04***
p-valor		0.84	0.94	0.15	0.00		0.53	0.74	0.45	0.01
Painel B: Alunos Negros										
Coefficiente	0.10	0.00	0.00	0.00	0.01***	0.10	0.00	-0.01	-0.01	0.00
p-valor		0.55	0.63	0.62	0.01		0.86	0.43	0.18	0.89
Painel C: Pai - Ensino Fundamental										
Coefficiente	0.20	0.01	0.01	0.01	0.03***	0.23	-0.01	0.00	0.01	0.03***
p-valor		0.11	0.06	0.06	0.00		0.29	0.92	0.32	0.00
Painel D: Pai - Ensino Médio										
Coefficiente	0.19	0.00	-0.01	-0.03***	0.08*	0.16	-0.01	-0.02	0.03*	-0.09***
p-valor		0.99	0.06	0.00	0.00		0.30	0.09	0.00	0.00
Painel E: Pai - Ensino Superior										
Coefficiente	0.02	0.00	0.00	-0.004**	0.02*	0.02	0.00	0.00	-0.004**	-0.02***
p-valor		0.50	0.50	0.03	0.00		0.43	0.22	0.05	0.00
Painel F: N° de Alunos (Iniciais/Finais)										
Coefficiente	329.18	-5.32	-19.56	-34.8**	-73.1***	376.53	-52.31	-38.18	-78.4***	-153.0***
p-valor		0.85	0.34	0.04	0.00		0.20	0.25	0.01	0.00
Painel G: Número de Funcionários										
Coefficiente	44.07	-4.30	-5.4**	-7.2***	-12.2***	56.93	-4.12	-6.16	-12.1***	-22.6***
p-valor		0.12	0.01	0.00	0.00		0.38	0.08	0.00	0.00
Painel H: Número de Salas Existentes										

Coefficiente	10.98	-0.90	-1.16**	-1.58***	-2.4***	11.61	-1.87**	-1.86**	-2.79***	-4.6***
p-valor		0.17	0.02	0.00	0.00		0.03	0.01	0.00	0.00
Painel I: Número de Computadores										
Coefficiente	20.14	0.62	0.65	0.35	-1.16	18.94	-0.79	-1.56	-1.85	-5.1***
p-valor		0.75	0.73	0.79	0.38		0.65	0.20	0.08	0.00
Painel J: Biblioteca										
Coefficiente	0.22	0.07	0.01	0.01	0.00	0.17	0.00	0.00	0.00	0.00
p-valor		0.20	0.76	0.69	0.98		1.00	0.96	0.95	0.91
Painel K: Refeitório										
Coefficiente	0.51	0.05	0.04	0.02	0.01	0.50	0.13	0.08	0.04	0.02
p-valor		0.45	0.43	0.59	0.80		0.23	0.29	0.57	0.71
Painel L: Sala Professores										
Coefficiente	0.95	0.01	-0.02	-0.03	-0.05***	0.97	0.02	-0.05	-0.03	-0.08***
p-valor		0.88	0.42	0.10	0.00		0.66	0.11	0.17	0.00
Painel M: Laboratório de Informática										
Coefficiente	0.87	-0.01	-0.02	-0.04	-0.03**	0.92	0.08	-0.01	-0.02	-0.04
p-valor		0.81	0.64	0.23	0.02		0.21	0.75	0.55	0.06
Painel N: Laboratório de Ciências										
Coefficiente	0.02	0.01	0.01	-0.02	-0.05**	0.12	0.05	0.02	-0.05	-0.24***
p-valor		0.66	0.54	0.38	0.02		0.46	0.72	0.33	0.00
Painel O: Quadra de Esportes										
Coefficiente	0.80	0.04	-0.02	-0.02	-0.04	0.90	0.11	0.06	0.02	-0.03
p-valor		0.48	0.60	0.50	0.21		0.08	0.15	0.57	0.15

*** Significativo a 1%

** Significativo a 5%

Tabela A.5: Balanceamento do Escore de Propensão (método do vizinho mais próximo) - Anos Iniciais e Anos Finais

Variável	Pareamento	Anos Iniciais			Anos Finais		
		Escolas Tratadas	Escolas Controle	p>t	Escolas Tratadas	Escolas Controle	p>t
Características Acadêmicas							
Ideb Iniciais/Finais - 2011	Antes	5.13	5.53	0.00	4.04	4.49	0.00
	Depois	5.16	5.16	1.00	4.05	4.03	0.74
Taxa de Abandono	Antes	0.34	0.25	0.04	2.42	1.52	0.00
	Depois	0.34	0.27	0.39	2.28	2.42	0.68
Taxa de Aprovação	Antes	94.86	95.92	0.00	88.98	91.84	0.00
	Depois	94.70	95.09	0.38	88.78	88.10	0.38
Taxa de Distorção Idade-Série	Antes	6.02	5.42	0.08	16.20	12.86	0.00
	Depois	6.03	5.97	0.92	17.04	17.37	0.70
Características Socioeconômicas							
% de Alunos Brancos (autodeclaração)	Antes	0.49	0.53	0.00	0.44	0.49	0.00
	Depois	0.49	0.48	0.43	0.44	0.43	0.20
% de Alunos Negros (autodeclaração)	Antes	0.08	0.08	0.48	0.10	0.09	0.00
	Depois	0.09	0.09	0.68	0.10	0.11	0.31
% de Pais com Ensino Fundamental	Antes	0.19	0.19	0.35	0.20	0.21	0.02
	Depois	0.19	0.19	0.62	0.20	0.20	0.43
% de Pais com Ensino Médio	Antes	0.19	0.22	0.00	0.17	0.21	0.00
	Depois	0.19	0.19	0.84	0.17	0.17	0.92
% de Pais com Ensino Superior	Antes	0.02	0.03	0.00	0.03	0.04	0.00
	Depois	0.02	0.02	0.71	0.03	0.03	0.78
% de Mães com Ensino Fundamental	Antes	0.21	0.20	0.18	0.23	0.23	0.56
	Depois	0.20	0.20	1.00	0.22	0.22	0.34
% de Mães com Ensino Médio	Antes	0.22	0.26	0.00	0.21	0.23	0.00
	Depois	0.22	0.22	0.61	0.20	0.21	0.96
% de Mães com Ensino Superior	Antes	0.03	0.04	0.00	0.03	0.05	0.00
	Depois	0.03	0.02	0.73	0.03	0.03	0.57
% de Pais Empregados	Antes	0.73	0.77	0.00	0.72	0.76	0.00
	Depois	0.73	0.73	0.92	0.72	0.72	0.97
% de Pais com Emprego Temporário	Antes	0.05	0.04	0.00	0.04	0.03	0.00
	Depois	0.05	0.05	0.81	0.04	0.04	0.95

% de Pais Desempregados	Antes	0.06	0.05	0.00	0.05	0.04	0.00
	Depois	0.06	0.06	0.54	0.04	0.04	0.48
% de Mães Empregadas	Antes	0.51	0.56	0.00	0.55	0.59	0.00
	Depois	0.51	0.51	0.85	0.55	0.55	0.62
% de Mães com Emprego Temporário	Antes	0.06	0.05	0.02	0.05	0.04	0.04
	Depois	0.06	0.06	0.93	0.05	0.05	0.68
% de Mães Desempregadas	Antes	0.23	0.20	0.00	0.18	0.17	0.00
	Depois	0.23	0.23	0.98	0.18	0.18	0.70
% de Famílias com Renda até 850,00	Antes	0.30	0.25	0.00	0.25	0.21	0.00
	Depois	0.30	0.30	0.86	0.25	0.25	0.59
% de Famílias com Renda entre 1276 e 2125	Antes	0.16	0.19	0.00	0.17	0.21	0.00
	Depois	0.15	0.15	0.48	0.17	0.17	0.77
% de Famílias com Renda maior que 2126	Antes	0.07	0.10	0.00	0.08	0.12	0.00
	Depois	0.07	0.07	0.98	0.08	0.08	0.76
% de Estudantes PBF	Antes	0.34	0.26	0.00	0.28	0.23	0.00
	Depois	0.36	0.35	0.92	0.28	0.27	0.68

Características Físicas e de Organização

Escola Estadual	Antes	0.11	0.36	0.00	0.80	0.81	0.56
	Depois	0.10	0.10	1.00	0.81	0.85	0.30
Escola Urbana	Antes	0.90	0.93	0.08	0.96	0.95	0.59
	Depois	0.91	0.90	0.70	0.95	0.98	0.24
Biblioteca	Antes	0.21	0.21	0.95	0.18	0.13	0.02
	Depois	0.24	0.24	1.00	0.16	0.12	0.34
Quadra de Esportes coberta	Antes	0.52	0.47	0.20	0.81	0.77	0.11
	Depois	0.56	0.55	0.91	0.82	0.83	0.77
Quadra de Esportes descoberta	Antes	0.30	0.34	0.25	0.30	0.38	0.01
	Depois	0.28	0.28	0.90	0.33	0.29	0.47
Número de Salas	Antes	11.44	11.26	0.66	13.91	12.99	0.01
	Depois	11.70	11.51	0.74	14.04	13.85	0.77
Número de Computadores	Antes	18.81	19.91	0.65	21.73	22.75	0.67
	Depois	18.34	18.20	0.91	21.51	20.75	0.53
Laboratório de Ciências	Antes	0.10	0.07	0.14	0.29	0.30	0.69
	Depois	0.12	0.12	0.87	0.34	0.33	0.91
Laboratório de Informática	Antes	0.81	0.79	0.56	0.95	0.94	0.41
	Depois	0.86	0.87	0.75	0.96	0.96	0.78

Sala de Diretoria	Antes	0.94	0.94	0.93	0.98	0.97	0.79
	Depois	0.94	0.96	0.43	0.98	0.99	0.41
Sala dos Professores	Antes	0.93	0.94	0.61	0.98	0.97	0.79
	Depois	0.94	0.97	0.19	0.98	0.98	1.00
Número de Funcionários	Antes	45.71	44.55	0.50	67.32	63.79	0.05
	Depois	46.20	46.53	0.90	69.12	68.45	0.82
Número de Alunos - Anos Iniciais	Antes	377.17	361.93	0.36	53.34	81.18	0.01
	Depois	387.76	403.87	0.50	50.87	53.96	0.84
Número de Alunos - Anos Finais	Antes	89.22	100.77	0.41	516.82	456.63	0.00
	Depois	83.06	76.39	0.74	515.83	506.84	0.76
Anos Iniciais/Finais	Antes	0.25	0.29	0.23	0.19	0.31	0.00
	Depois	0.21	0.20	0.79	0.18	0.20	0.68
Ensino Médio	Antes	0.08	0.12	0.06	0.76	0.72	0.26
	Depois	0.07	0.05	0.48	0.77	0.79	0.69
Educação Infantil	Antes	0.32	0.22	0.00	0.01	0.03	0.04
	Depois	0.29	0.33	0.40	0.01	0.00	0.32
<hr/>							
Características dos Municípios							
Pib per Capita	Antes	22615	29692	0.00	29922	30260	0.78
	Depois	22924	26289	0.02	30002	32403	0.22
População	Antes	320000	1600000	0.00	1400000	1600000	0.42
	Depois	340000	340000	0.99	1800000	2000000	0.58
IDHM	Antes	0.76	0.77	0.00	0.76	0.77	0.03
	Depois	0.76	0.77	0.22	0.77	0.77	0.26
Total de Observações:		163	3,564		163	3,123	

Fonte: Elaboração própria a partir dos dados pesquisados.

Ensaio III

3 ENSAIO III: Tecnologia e Desempenho Escolar: Uma Avaliação de Impacto do Programa Um Computador por Aluno

Resumo

Este trabalho analisa os efeitos da inserção de tecnologias de informação e comunicação (TIC) no ambiente escolar no desempenho acadêmico de estudantes. Para isso, realizamos a primeira avaliação de impacto rigorosa do Programa Um Computador por Aluno – Prouca, que forneceu um *laptop* para cada estudante, além de 120 horas de treinamento para docentes, em mais de 300 escolas brasileiras em 2010. Os resultados indicam que, embora o programa tenha impactado positivamente a razão computador por aluno e a utilização de computadores e da internet para fins pedagógicos, não há evidência de efeitos significativos no desempenho acadêmico dos estudantes em testes padronizados de matemática e português e nas taxas de abandono. Esses resultados mostram-se coerentes com a literatura recente de avaliação de impacto de programas 1:1 em países em desenvolvimento.

Palavras-chave: Educação; Tecnologia; Desempenho Escolar; Avaliação de Impacto.

3.1 Introdução

Nas décadas de 1980 e 1990, o Estado brasileiro implementou política extremamente bem-sucedida de ampliação do acesso à educação básica. A taxa de matrícula líquida no Ensino Fundamental, que era de 64% em 1980, alcança 86% no início da década seguinte, 95% em 2001 e 97% em 2007, quando se considera o ensino básico no país praticamente universalizado⁶⁰ (Rigotto & Souza, 2006). Por outro lado, pouco se tem avançado no sentido de aumentar a qualidade da educação, apesar do crescimento substantivo do investimento em educação nas últimas décadas. Esse é um cenário que caracteriza não apenas o Brasil, mas diversos países em desenvolvimento⁶¹.

Nesse contexto, vem ganhando relevância o debate acerca da utilização de tecnologias da informação e comunicação (TICs) no ambiente escolar como forma de aumentar a produtividade da educação. Os computadores e a internet já podem ser considerados importantes componentes da educação moderna. A título de exemplo, todas as salas de aulas de escolas públicas dos Estados Unidos possuem, atualmente, computadores com acesso à internet (U.S. Department of Education, 2013)⁶². O mesmo ocorre na Europa, onde a maioria dos países também possui altos índices de acesso a computadores e à internet nas escolas (European Commission, 2013). Países da América Latina, embora não tenham alcançado patamares semelhantes aos dos países desenvolvidos, têm se esforçado nesse sentido. Segundo UNESCO (2012), trinta e um países da América Latina e do Caribe tinham políticas de inserção das TICs no ambiente escolar em 2012.

O grande entusiasmo com o qual atores políticos e gestores públicos encaram a inserção de TICs no ambiente escolar não parece encontrar, contudo, respaldo na literatura sobre o tema. Como destacam Bulman & Fairlie (2016), em recente revisão da literatura, a evidência de estudos empiricamente rigorosos sobre a utilização da tecnologia

⁶⁰ Dados da Pesquisa Nacional por Amostras de Domicílio (PNAD) de 2009 mostram que a taxa de cobertura escolar para crianças de 7 a 14 anos atingiu 98% e para crianças de 6 a 14 anos, 97%.

⁶¹ Muralidharan et al.(2016), por exemplo, descrevem um cenário bastante semelhante para a Índia.

⁶² Como destacam Grimes & Warschauer (2008), a razão computador por aluno nos Estados Unidos caiu de aproximadamente 168:1 em 1983 para cerca de 3,8:1 em 2005.

na educação é mista, com uma predominância de resultados nulos, principalmente em países em desenvolvimento.

Este trabalho pretende contribuir para esta literatura, por meio da realização da primeira avaliação de impacto rigorosa do Programa Um Computador por Aluno – Prouca. O Prouca se insere na categoria de políticas 1:1, caracterizadas pela distribuição de um *laptop* para cada estudante e professor da escola. Em 2010, o programa distribuiu cerca de 150 mil *laptops* a 367 escolas públicas brasileiras. Além disso, o programa forneceu cerca de 150-180 horas de treinamento para os professores das escolas beneficiárias.

Chama a atenção o fato de que não se tenha conduzido, até o presente momento, nenhum estudo rigoroso sobre o impacto do Prouca no desempenho acadêmico dos estudantes. Ainda que não sejam poucos os estudos realizados com o intuito de analisar o programa, todos os trabalhos revisados apoiam-se em metodologias qualitativas, em grande medida baseados em estudos de caso e questionários. Trata-se do que Penuel (2006) considera “estudos de implementação” e não de resultados. Assim, nenhum dos trabalhos produzidos permite uma interpretação causal dos impactos do programa⁶³.

Avaliamos o impacto do Prouca no desempenho acadêmico dos estudantes em português e matemática (medido por meio dos exames do Sistema de Avaliação da Educação Básica - SAEB), nas taxas de abandono e na utilização do computador e da internet por docentes para fins pedagógicos. Em linha com a literatura sobre o tema, a avaliação mostrou que o programa, embora tenha impactado positivamente o número de computadores nas escolas tratadas e a utilização do computador e da internet pelos docentes, não apresentou impacto significativo no desempenho acadêmico dos estudantes e nas taxas de abandono. Trata-se de resultado consistente com diversas outras avaliações de impacto de programas 1:1, particularmente em países em desenvolvimento, como veremos na seção seguinte.

O restante deste artigo está organizado em 6 seções, além desta Introdução. Na seção 3.2, revisaremos brevemente a literatura acerca da utilização de TICs no ambiente escolar, com foco em avaliações de impacto empíricas. Na seção 3.3, apresentaremos as principais características do Programa Um Computador por Aluno e de sua

⁶³ Para revisões de literatura sobre artigos, dissertações e teses produzidas sobre o Prouca, vide Santos (2014), Gomes (2015) e Andriola et al. (2017).

implementação. A seção 3.4 discutirá a estratégia empírica e os modelos econométricos utilizados. Na seção 3.5, apresentaremos as bases de dados e na seção 3.6 analisaremos os principais resultados e exploraremos possíveis efeitos heterogêneos. A seção 3.7 traz a conclusão deste artigo.

3.2 Revisão da Literatura

A utilização da tecnologia pelas escolas pode ocorrer de diversas maneiras. Para melhorar compreender o resultado dessas iniciativas, convém classificarmos as políticas de “informatização” das escolas em dois tipos : i) investimento em TICs (*hardwares*), tal como a compra de computadores e a instalação de conexão para acesso à internet; ii) investimentos em *softwares* específicos para o ensino auxiliado pelo computador – *computer aided instruction* – CAI (Bulman & Fairlie, 2016).

Intervenções baseadas em *hardware*, em sua maioria, não apresentam impacto positivo no aprendizado. Angrist & Lavy (2002), em uma das primeiras avaliações rigorosas de impacto sobre o tema, analisaram um programa israelense que forneceu computadores e treinamentos para professores em escolas de ensino fundamental e médio. Empregando diferentes técnicas de estimação, os autores encontram um impacto negativo e marginalmente significativo nas notas de matemática para alunos da 4^a série e insignificante para as demais séries e disciplinas. Em linha com Angrist & Lavy (2002), Leuven et al. (2007) avaliaram, por meio de uma regressão com descontinuidade, política holandesa que forneceu financiamento para aquisição de computadores em escolas com mais de 70% de estudantes em condições de desvantagem. Também nesse caso, os autores encontraram efeitos negativos, mas insignificantes no desempenho em testes escolares. Em ambos os casos, os autores conjecturam que esse efeito pode advir do fato de o ensino por computadores ser menos efetivo do que o ensino tradicional.

Goolsbee & Guryan (2006), por sua vez, avaliaram os efeitos de um programa de subsídios à instalação e utilização de internet na Califórnia (E-Rate Program). Embora o programa tenha impactado na expansão do acesso à internet nas escolas tratadas, os autores não encontraram evidência de impacto significativo do programa no desempenho acadêmicos dos estudantes. Sprietsma (2012), por sua vez, avaliou o efeito do aumento da disponibilidade de computadores e da internet na escola, declarados no Censo Escolar, no desempenho acadêmicos de alunos da 8^a série no Brasil. As estimativas mostraram que o aumento da disponibilidade de computadores teve efeito negativo no desempenho escolar, enquanto o aumento da disponibilidade da internet teve efeito positivo.

Programas do tipo 1:1, que se baseiam na distribuição de *laptops* e *tablets* para professores e alunos, são exemplos de intervenções baseadas em *hardware*. Nos últimos dez anos, observou-se um crescimento significativo de programas do tipo em todo mundo, particularmente em países em desenvolvimento. Este crescimento foi diretamente influenciado pela iniciativa Um Computador por Criança (*One Laptop per Child - OLPC*), lançada em 2005.

Após poucos anos, o projeto logrou alcance mundial, incluindo diversos países na América Latina, África, Ásia e América do Sul⁶⁴. Países como o Peru (580 mil *laptops*), Uruguai (420 mil) e Ruanda (110 mil) são somente alguns exemplos de adoção de larga escala do projeto⁶⁵. Em 2011, mais de 2 milhões de *laptops* OLPC já haviam sido distribuídos em mais de 40 países⁶⁶. Infelizmente, contudo, a implementação dessas iniciativas não foi precedida de estudos empíricos que comprovassem a eficácia da estratégia (Nugroho & Lonsdale, 2010). Avaliações de impacto de programas 1:1 mostram evidências inconclusivas, com uma prevalência de resultados nulos, principalmente em países em desenvolvimento

Cristia et al (2012), por exemplo, avaliaram a implantação do OLPC em 319 escolas primárias rurais no Peru. Os autores lançaram mão de uma avaliação aleatória e os resultados indicam que, apesar de o programa ter aumentado a razão de computador por aluno de 0,12 para 1,18 nas escolas de tratamento, essa expansão não teve qualquer efeito no número de matrículas e no desempenho acadêmico em testes de Linguagem e de Matemática. Por outro lado, os autores encontraram impacto positivo em testes cognitivos.

Avaliações de impacto de programas de distribuição de *laptops* na Colômbia, no Uruguai e na Argentina e em Honduras chegaram a conclusões semelhantes. Linden & Barrera-Osorio, (2009) utilizaram um desenho aleatório para avaliar o programa *Computadores para Educar* em 97 escolas colombianas e não encontraram efeito do programa no desempenho acadêmico dos alunos. Segundo os autores, o resultado se deve, ao menos em parte, ao fato de os docentes não terem sido capazes de implementar novas

⁶⁴ Vide Nugroho & Lonsdale (2010).

⁶⁵ Vide Hansen et al. (2012).

⁶⁶Vide Verma (2011). Há, ainda, iniciativas em diversos outros países, tais como Gaza, Afeganistão, Haiti, Etiópia e Mongólia. Para mais exemplos, vide <http://one.laptop.org/about/contries>.

técnicas em sala de aula, apesar dos treinamentos recebidos. Já De Melo et al. (2014b) utilizaram o método de diferenças-em-diferenças para avaliar a implantação do OLPC no Uruguai⁶⁷ e tampouco encontraram efeitos positivos nas notas dos estudantes em leitura e matemática. Alderete & Formichella (2017) avaliaram, ainda, o programa *Conectar Igualdad*, que tem como objetivo prover infraestrutura tecnológica e computadores na razão 1:1 pelos estudantes das escolas (em novembro de 2014, mais de 4,5 milhões e meio de computadores já haviam sido distribuídos). Os autores empregaram a técnica de diferenças em diferenças com escore de propensão às notas dos estudantes no *Programme for International Student Assessment* – PISA em 2012 e encontraram um efeito significativo do programa no desempenho acadêmico dos estudantes. Contudo, esse efeito foi muito pequeno em magnitude, o que leva os autores a afirmarem que não se verifica uma “diferença qualitativa” no desempenho acadêmico dos estudantes. Finalmente, Bando et al. (2017) avaliaram, por meio de um desenho aleatório em 271 escolas, o impacto de se substituir livros didáticos por computadores em Honduras, no âmbito do programa Educatracho. Os autores mostram que, ao final de um ano, o programa não impactou de forma significativa o aprendizado dos alunos em matemática e linguagem.

Trabalhos realizados em países asiáticos e africanos igualmente, em sua maioria, não encontraram impacto de programas 1:1 no desempenho acadêmico dos estudantes. Sharma, Sharma & Uttam (2014) estimaram por meio de diferenças-em-diferenças o efeito de um pequeno piloto conduzido por uma ONG em 26 escolas do Nepal. Os autores não encontraram efeitos significativos em matemática e encontraram efeitos negativos em linguagem. Mo et al. (2013) avaliaram o impacto de uma intervenção inspirada no OLPC em 13 escolas de migrantes em Pequim. Os autores encontraram efeitos positivos em matemática (somente significativo a 10%) e nenhum efeito em leitura. Hansen et al. (2012), por sua vez, avaliaram a distribuição de 4.375 *laptops* OLPC na Etiópia. Os autores tampouco encontraram impacto significativo do programa nas notas médias de linguagem e matemática. Por outro lado, encontraram efeito significativo e positivo em testes relacionados a capacidade de “raciocínio abstrato”.

Mesmo em países desenvolvidos, onde a implementação de programas do tipo já ocorre há mais tempo, não há clareza na literatura sobre seus reais impactos. Penuel

⁶⁷ O Uruguai é o primeiro país a implementar a iniciativa OLPC nacionalmente (De Melo et al., 2014b).

(2006) avaliou cerca de 30 artigos sobre iniciativas 1:1. Chama a atenção que apenas 4 desses estudos tenham utilizado desenhos de pesquisa quase-experimentais rigorosos. Em nenhum desses, foi possível detectar impacto do programa no desempenho acadêmico. Nugroho & Lonsdale (2010) chegaram à conclusão semelhante: que a avaliação de programas OLPC, em geral, baseia-se em evidências anedóticas e que a maior parte dos trabalhos tratavam de aspectos técnicos e pedagógicos relacionados à implementação do programa, mas não avaliavam seus impactos.

Isso não significa, contudo, que não haja na literatura exemplos de impactos positivos de programas de inserção de *laptops* em escolas. Alguns estudos encontraram efeitos positivos e significativos desses programas no “engajamento dos estudantes”⁶⁸, em habilidades relacionadas à tecnologia⁶⁹, no desempenho acadêmico em linguagem e em matemática⁷⁰.

Por outro lado, políticas que se baseiam em *softwares* específicos para o ensino auxiliado pelo computador (*computer aided instructon* - CAI) têm apresentado evidências mais consistentes de impactos positivos. Esses programas, em geral, baseiam-se em metodologias de ensino que permite o acompanhamento e a progressão individualizados como forma de se lidar com as diferenças entre os estudantes.

Um dos estudos mais conhecidos do impacto de uma intervenção baseada em CAI foi conduzido por Banerjee et al. (2007), que avaliaram um programa indiano no qual instrutores auxiliavam estudantes na utilização de softwares educacionais durante duas horas por semana (sendo uma delas fora do horário tradicional da escola). Os autores encontraram um efeito positivo e estatisticamente significativo nas notas de matemática a curto prazo. O impacto, contudo, desapareceu ao longo dos anos. Em linha com esses resultados, Muralidharan et al. (2016) avaliaram, com base em um experimento aleatório, um programa de CAI chamado *Mindspark* implementado em estudante da 6^a a 9^a série em Nova Deli, Índia. Os autores encontraram um impacto positivo e significativo tanto em português quanto em linguagem. Os efeitos permaneceram mesmo quando considerados os diferentes níveis de desempenho inicial dos estudantes, mas os ganhos relativos foram maiores para alunos com pior desempenho acadêmico inicial.

⁶⁸ Vide, por exemplo, Bebell & Kay (2010), Trimmel & Bachmann (2004) e Kozma et al. (2004)

⁶⁹ Vide, por exemplo, Shapley (2009), Trimmel & Bachmann (2004) e Kozma et al. (2004)

⁷⁰ Vide, por exemplo, Machin, McNally, & Silva (2006), Suhr et al. (2010) e Shapley (2009)

Por outro lado, Linden (2008), em duas avaliações de impacto, encontrou efeitos distintos para programas realizados dentro da escola (ou seja, que substituíam, em parte, o ensino tradicional) e fora da escola (que complementavam o ensino tradicional). No primeiro caso, o autor encontrou efeitos negativos e significativos, de magnitude considerável; no segundo caso, o autor encontra efeitos positivos, mas estatisticamente não significativos. De forma semelhante, Spiezia (2011) mostra que o uso do computador tem efeitos positivos no desempenho dos estudantes no PISA, contudo, o aumento da nota advém, principalmente, do uso do computador em casa; os efeitos do uso do computador na escola, segundo o autor, é bastante pequeno ou inexistente.

Carrillo et al. (2016), por sua vez, avaliaram a implementação do Programa Más Tecnología, no município de Guayaquil, no Equador, por meio de um experimento aleatório. O programa foi aleatorizado a nível da escola e previa o fornecimento de 4 computadores por escola e de, pelo menos, 3 horas por semana de instrução personalizada em português e matemática para cada aluno. Os autores encontram efeitos positivos e significativos do programa em matemática e efeitos negativos, porém não significativos, em linguagem.

A compreensão sobre como a inserção da tecnologia nas salas de aula afeta o ensino e o aprendizado é de fundamental importância. Como vimos, diversos países vêm investindo somas significativas de recursos em programas que visam ampliar o acesso a computadores e a internet em sala de aula. Não encontramos evidências sólidas, contudo, de que as TICs sejam, por si só, insumos importantes na função de produção da educação. Pelo contrário, a evidência existente, principalmente no tocante a programas focados em *hardware*, é que essas intervenções tenham impacto nulo no desempenho acadêmico dos estudantes.

É importante ressaltar que, particularmente em um contexto caracterizado pela restrição de orçamento e da quantidade de tempo instrucional, não basta que a tecnologia tenha impacto positivo no ensino e no aprendizado; ela deve ter um impacto superior a insumos tradicionais. Não se pode perder de vista que o financiamento de programas de informatização das escolas ocorre em detrimento do investimento desses mesmos recursos em insumos tradicionais, como livros, professores, monitores e etc. Da mesma forma, o tempo dedicado ao ensino por meio de computador ocorre, na maioria dos casos, em detrimento da utilização desse mesmo tempo em técnicas tradicionais de ensino.

A evidência inconclusiva da literatura a respeito do impacto dessas políticas no desempenho acadêmico dos estudantes deve ser, assim, levada em consideração por

gestores públicos no desenho de programas que visem aumentar o acesso a computadores e a internet nas escolas. Os resultados dependem, em larga medida, da tecnologia adotada e em como essa tecnologia é integrada no dia-a-dia escolar. A revisão da literatura de programas focados em *softwares*, que adotam uma estratégia mais bem definida sobre como os computadores serão utilizados por docentes e estudantes, apresentam resultados superiores a programas que previram, simplesmente, o aumento do número de computadores ou do acesso à internet nas escolas, sem a devida atenção aos detalhes sobre como essa tecnologia seria inserida no processo de aprendizado. Ainda no caso de programas baseados em *softwares*, contudo, a evidência pode ser considerada mista, com várias intervenções apresentando resultados nulos, o que indica a necessidade de atenção aos detalhes na análise de cada uma dessas intervenções.

O Prouca é claramente um exemplo de programa baseado em *hardware*. Nesse caso, como veremos, muita atenção foi dada à disponibilização de *laptops* para os estudantes, enquanto a forma de inserção desses aparelhos no dia-a-dia escolar foi relegada a um segundo plano - ainda que não tenha sido de todo desconsiderada. Não surpreende, portanto, que o investimento de quase R\$ 100 milhões realizado no programa não tenha tido qualquer impacto no desempenho acadêmico dos estudantes e nas taxas de abandono.

3.3 O Programa Um Computador por Aluno - Prouca

A iniciativa OLPC foi apresentada ao Governo Brasileiro, em 2005, no Fórum Econômico Mundial⁷¹. Dois anos depois, em 2007, teve início a fase piloto⁷² da versão nacional do projeto em cinco cidades brasileiras: Pirai (RJ), Porto Alegre (RS), Palmas (TO), Brasília (DF) e São Paulo (SP). Em cada uma dessas cidades, foi escolhida uma escola⁷³ e todos os seus alunos receberam diferentes modelos de *laptops* educacionais de

⁷¹ Em meados de 2005, os professores Nicholas Negroponte, Seymour Papert e Mary Lou Jepsen vieram ao Brasil apresentar o projeto a membros do Governo Federal.

⁷² Há diversas referências, em documentos oficiais, a essa fase como “pré-piloto” e ao Prouca, de fato, como “piloto”. Neste trabalho, nos referimos a essa primeira experiência como “piloto”, uma vez que consideramos o Prouca o programa implementado.

⁷³ Foram elas: Escola Municipal Ernani Bruno no município de São Paulo (SP), Escola Estadual Luciana de Abreu em Porto Alegre (RS), Colégio Estadual Dom Alano Marie Du Noday em Palmas (TO), CIEP

baixo custo, cedido gratuitamente por três fabricantes (OLCP – modelo *XO*, Intel – Modelo *ClassMate* e Encore – modelo *Mobilis*).

Ainda nesse ano, foi criada a equipe UCA, por meio de Portaria do Ministério da Educação, com o objetivo de estruturar propostas de implementação, pesquisa, monitoramento e avaliação do programa. Essa equipe elaborou, em março de 2007, os princípios norteadores para o uso pedagógico do laptop no país (SEED/MEC, 2007).

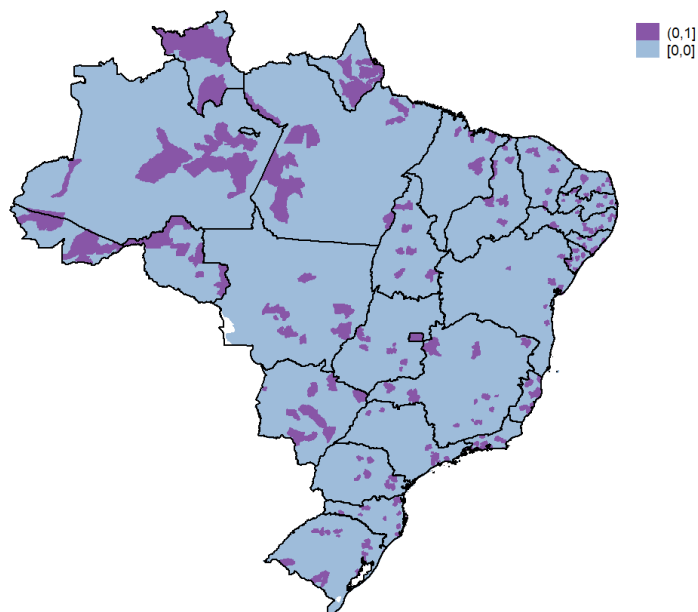
O Programa foi implementado em 2010 e abrangeu mais de 350 escolas, em todos os estados da federação. No total, foram distribuídos aproximadamente 150 mil *laptops* (cerca de 400 *laptops* por escola)⁷⁴. Em média, foram contempladas 10 escolas por estado (exceto no caso de estado com municípios no Programa UCA-Total⁷⁵). Como se nota na Figura 14, o programa foi bem distribuído entre as diversas regiões e estados do país.

Municipal Prof. Rosa Conceição Guedes em Piraí (RJ) e Centro de Ensino Fundamental no 1 do Planalto em Brasília (DF).

⁷⁴ O laptop do programa foi selecionado por meio de concorrência pública. O consórcio CCE/Digibras/Metasys foi o vencedor do pregão nº 107, de 2008. O computador vencedor possui configuração exclusiva para o programa e atende aos requisitos funcionais do programa: possui bateria com autonomia mínima de 3 horas, peso de até 1,5kg, com 1 GB (um gigabyte) de memória RAM e armazenamento de 8 GB. O laptop veio equipado com sistema operacional Linux (software livre). Além disso, sua configuração padrão vinha com alguns utilitários, que incluíam: calculadora científica, aplicativo de textos simples, programas de aprendizado infantil, arte e música, digitação para crianças e matemática. O modelo do laptop despertou algumas críticas, em virtude do tamanho da tela (muito pequena), da capacidade limitada de armazenamento e da bateria (Pinheiro, Rosa, & Bonilla, 2012). O custo por laptop foi de R\$ 550,00 (ou US\$ 316,00), o que supera significativamente a expectativa inicial do OLPC de distribuir computadores ao custo de US\$ 100,00 por aparelho. As entregas dos computadores ocorreram de maio a novembro de 2010.

⁷⁵ O Programa teve, ainda, uma versão intitulada UCA-Total, em que seis municípios foram selecionados: Barra dos Coqueiros (SE), Caetés (PE), Santa Cecília do Pavão (PR), São João da Ponta (PA), Terrenos (MS) e Tiradentes (MG). Nesses municípios, todas as escolas – municipais e estaduais – foram contempladas com laptops do Prouca. Fora essa característica, esse experimento em nada difere do Prouca.

Figura 14– Distribuição Nacional do Programa Um Computador por Aluno



Fonte: Elaboração própria com base em dados fornecidos pela Secretaria de Educação Básica do Ministério da Educação.

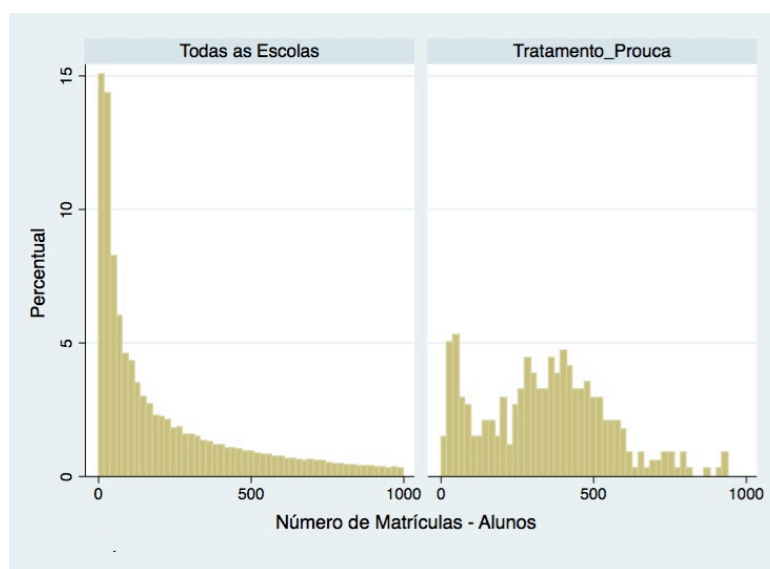
A seleção das escolas participantes foi feita pelas Secretarias de Educação (estadual ou municipal). Essa seleção deveria respeitar 5 critérios (BRASIL, 2013)

- i. Número de alunos e professores – as escolas deveriam ter em torno de 500 alunos e professores;
- ii. Estrutura da escola – as escolas deveriam possuir, obrigatoriamente, energia elétrica para carregamento dos *laptops* e armários para armazenamento;
- iii. Localização das escolas – deveriam ser selecionadas escolas preferencialmente próximas a Núcleos de Tecnologias Educacionais – NTE – ou similares, Instituições de Educação Superior Públicas ou Escolas Técnicas Federais. Além disso, pelo menos uma das escolas deveria estar localizada na capital do estado e uma na zona rural);
- iv. Assinatura de termo de adesão – As Secretarias de Educação Estaduais ou Municipais de cada uma das escolas deveriam “aderir” ao projeto por meio de envio de ofício ao MEC e assinatura de Termo de Adesão

- v. Anuência do corpo docente - para cada escola, a Secretaria de Educação deveria enviar ao MEC um ofício onde o diretor da escola, com anuência do corpo docente, aprova a participação no projeto.

Como mostra a Figura 15, apesar da existência do critério formal de 500 alunos e professores, esse critério não foi respeitado na seleção final das escolas. Cerca de 25% das escolas selecionadas tinham menos de 200 alunos e 10% delas mais de 600.

Figura 15 – Histograma do número de alunos por escola



Fonte: Elaboração própria a partir dos dados pesquisados.

O Prouca tinha como objetivo principal promover *uma nova forma de utilização das tecnologias digitais nas escolas públicas, balizada pela necessidade de: i) melhoria de qualidade da educação; b) inclusão digital; c) inserção da cadeia produtiva brasileira no processo de fabricação e manutenção de equipamentos* (SEED/MEC, 2007).

Para alcançar esses objetivos, o programa não se limitou à distribuição dos *laptops*, mas também previu um processo de capacitação de recursos humanos (professores e alunos-monitores) e de acompanhamento e avaliação do programa, que permitisse *valorar o impacto do laptop educacional conectado no processo educacional em termos pedagógicos e organizacionais para a melhoria na qualidade do ensino e da aprendizagem e inclusão digital da comunidade escolar* (UFC, 2010). Nesse sentido, a intervenção realizada no âmbito do programa inclui tanto a distribuição dos *laptops* quanto um treinamento de 150-180 horas para os docentes das escolas contempladas pelo

programa, que teve início juntamente com a distribuição de computadores, em meados de 2010⁷⁶.

Para implementação das ações previstas no âmbito do Prouca, o Governo Federal adotou a seguinte repartição de competências: a Secretaria de Educação à distância do Ministério da Educação - MEC⁷⁷ e o Fundo Nacional de Desenvolvimento da Educação – FNDE forneceriam os computadores e coordenariam a formação dos professores e as redes municipais e estaduais seriam responsáveis por preparar a infraestrutura elétrica e o funcionamento da internet sem fio (Pretto, Coelho, & Almeida, 2012)⁷⁸.

Uma característica fundamental de programas 1:1 diz respeito à possibilidade de os estudantes levarem os *laptops* para casa. Nesse aspecto, apesar de a orientação oficial dar a entender a possibilidade de utilização do laptop em “ambientes dentro e fora da escola (SEED/MEC, 2007), a implementação do programa pode ser considerada bastante heterogênea, cabendo a cada escola e/ou secretaria de educação essa definição⁷⁹.

Finalmente, é importante destacar que parece ter havido uma espécie de “descontinuidade silente” do Programa. Embora não haja manifestação oficial do

⁷⁶ A formação dos professores foi realizada por profissionais do NTE, com orientação e acompanhamento das IES Globais e Locais, e foi estruturada em 5 módulos, com duração de 180 horas, sendo 150 deles obrigatórios (UFC, 2010): i) Módulo 1 (40 horas presenciais) – apropriação tecnológica e compreensão da proposta UCA; ii) Módulo 2 (30 horas à distância) – Recursos web 2.0; iii) Módulo 3 (40 horas à distância) – formação na escola proposta da ação; iv) Módulo 4 (40 horas presenciais) – análise e depuração da proposta de ação e elaboração do Projeto UCA da Escola; v) Módulo 5 - optativo (3 dias – integrada ao Módulo 4) – seminário para apresentação dos resultados, análises com base em documentos e nas apresentações do projeto UCA da Escola.

⁷⁷ Com a extinção da SEED/MEC, em 2011, o Programa foi alocado na Secretaria de Educação Básica (SEB) do Ministério.

⁷⁸ Adicionalmente, desempenham atividades importantes no Programa os seguintes atores: i) Grupo de Trabalho de Assessoramento Pedagógico do UCA (GT-Uca) – responsável pelo assessoramento pedagógico, a elaboração do documento básico do projeto e o acompanhamento e avaliação do pré-piloto e do piloto do Programa; ii) Instituições de Ensino Superior (IES) Globais – preparação da equipe de formação dos membros das IES locais; iii) Instituições de Ensino Superior (IES) Locais – formação de professores e gestores da escola; e os iv) Núcleos de Tecnologia Estadual (NTE) e Municipal (NTM) – formação de professores e gestores da escola

⁷⁹ Em avaliação qualitativa da implantação do Prouca em seis escolas no Pará, Costa (2013) cita que apenas uma das seis escolas permitia que o laptop fosse levado para casa. Santos (2014), ao avaliar a implementação em escolas de Goiânia, descreve cenário semelhante, assim como Lavinias & Veiga (2013) ao avaliarem a implementação do Programa Uca Total.

Governo Federal a respeito, não parece existir mais qualquer forma de auxílio - financeiro, técnico ou de qualquer outra forma - para escolas do programa. É simbólico dessa descontinuidade o fato de o portal do programa ter sido retirado do ar em 2015, sem qualquer tipo de comunicado oficial⁸⁰. Segundo Schneider (2012), o site do programa não vinha sendo atualizado desde dezembro de 2010, o que indica que o relativo “abandono” do programa pelo Governo Federal pode ter ocorrido muito antes de 2015, logo após a distribuição dos *laptops*. Como consequência, a base de documentos oficiais sobre o programa encontra-se, atualmente, dispersa e muitos documentos não encontram registros disponíveis para acesso público.

3.4 Estratégia Empírica

3.4.1. Métodos de Pareamento – *Propensity Score Matching* (PSM)

Na ausência de aleatorização, uma das formas mais comuns de tentar eliminar o viés de seleção é construindo um grupo controle que seja o mais parecido possível com o grupo de tratamento. Para isso, seleciona-se um conjunto de características observáveis e, a partir dessas características, tentamos encontrar, para cada unidade tratada, uma unidade não-tratada que apresente exatamente as mesmas características. Esses métodos são conhecidos como pareamento (*matching*). A intuição do método é que para cada unidade tratada existiria uma ou mais unidades de controle suficientemente parecidas com ela, de modo a representar o resultado que teria corrido caso a unidade tratada não tivesse sido submetida ao tratamento.

Note que quanto maior o número de variáveis de controle - principalmente se forem contínuas e/ou multidimensionais -, maior será a dificuldade em se encontrar uma escola no grupo de controle que possua as mesmas características de uma escola do grupo de tratamento. Essa dificuldade, contudo, foi contornada por meio do método de “pareamento por escore de propensão” (PSM) (Rosenbaum & Rubin, 1983)⁸¹.

⁸⁰ Vide o site: <http://www.uca.gov.br>. Tentativas de acesso realizada entre os meses de março e junho de 2017.

⁸¹ Para uma apresentação forma do método e completa do método, vide Rosenbaum & Rubin (1983), Dehejia & Wahba (1998) e (Becker & Ichino, 2002).

Para implementação do método, calculamos, inicialmente, a probabilidade de uma escola receber o tratamento com base em características observáveis (que serão detalhadas na seção seguinte) por meio de um modelo *probit*⁸². O escore de propensão é, então, definido por Rosenbaum & Rubin (1983) como a probabilidade condicional de se receber o tratamento, dada a existência de determinadas características pré-tratamento:

$$p(X) \equiv \Pr(D = 1|X) = E(D|X) \quad (1)$$

Em seguida, dividimos o grupo de escolas em dois grupos (tratamento e controle) e para cada escola no grupo tratamento buscaremos uma ou mais escolas no grupo controle com um escore de propensão bastante próximo. Em outras palavras, pareamos cada escola tratada com uma ou mais escolas no grupo de controle que apresentem características bastante similares⁸³. As unidades de controle que tenham um escore de propensão próximos aos de unidades de tratamento formarão o grupo de controle utilizado como contrafactual.

Como ensinam Becker & Ichino (2002), atendidas as hipóteses de identificação conhecidas como “ignorabilidade forte”⁸⁴, observações com o mesmo escore de propensão devem ter a mesma distribuição de características (observáveis e não observáveis) independentemente do *status* do tratamento. Ou seja, a designação do tratamento é

⁸² Por definição, o valor de $p(x)$ não é observado, mas estimado com base em um modelo estatístico de probabilidade. Como destacam Caliendo & Kopeinig (2008), há pouca orientação disponível sobre a forma funcional a ser usada para a estimação do escore de propensão. A princípio, pode-se empregar qualquer modelo de probabilidade padrão pode ser utilizado. A escolha em geral recai em modelos *probit* e *logit*, dada as limitações do modelo de probabilidade linear. Os autores destacam que, no caso de tratamentos binários, onde estima-se a probabilidade de participar *versus* de não participar, os modelos *logit* e *probit* em geral levam aos mesmos resultados. Na estimação adotada nesse trabalho, adotaremos o modelo *probit*.

⁸³ Parear, no caso, significa nada mais do que encontrar para cada unidade de tratamento um ou mais controles que tenham um escore de propensão igual ou bastante próximo.

⁸⁴ Para uma descrição detalhada dessas hipóteses, vide Rosenbaum & Rubin (1983). Uma das etapas fundamentais do artigo é demonstrar que, caso a atribuição do tratamento seja aleatória dentro das células definidas por X (ou seja, dentro de cada grupo com as mesmas características pré-tratamento), então também será aleatória dentro das células definidas pelos valores do escore de propensão (ou seja, dentro de cada grupo com o mesmo escore de propensão).

aleatória, dado um determinado escore de propensão. Sob a hipótese de independência, o pareamento com base em $p(x)$ elimina, portanto, o viés de seleção.

Atendida as hipóteses, o efeito médio do tratamento nos tratados (ATT) pode ser estimado como (Becker & Ichino, 2002):

$$\begin{aligned}
 ATT &= E\{Y_i^T - Y_i^C | D_i = 1\} & (4.2) \\
 &= E[E\{Y_i^T - Y_i^C | D_i = 1, p(X_i)\}] \\
 &= E[E\{Y_i^T | D_i = 1, p(X_i)\} - E\{Y_i^C | D_i = 0, p(X_i)\} | D_i = 1]
 \end{aligned}$$

onde a expectativa é sobre a distribuição de $p(X_i)|D_i=1$ e Y_i^T e Y_i^C são resultados potenciais nas duas situações contrafactuais de tratamento (Y_i^T) e não tratamento (Y_i^C).

Em outras palavras, podemos adotar como contrafactual do grupo de tratamento os resultados observados do grupo de controle pareado (ou seja, os resultados do conjunto de escolas não tratadas que possuem o mesmo escore de propensão que as escolas do grupo de tratamento). Note que o estimador do PSM é nada mais do que a diferença média nos resultados sobre o suporte comum, ponderados pela distribuição do escore de propensão dos participantes.

Note que, como $p(x)$ é uma variável contínua, quando utilizamos um grande número de controles, a probabilidade de encontrarmos duas observações com dois escores de propensão idênticos é, no limite, zero. Nesse contexto, existem vários métodos para superar esse problema e realizar o pareamento, dentre os quais se destacam: pareamento pelo vizinho mais próximo (*nearest-neighbor matching*), pelos 5 vizinhos mais próximos (*nearest 5-neighbor matching*), pareamento por raio (*radius matching*) e pareamento *kernel* (*Kernel matching*). De forma a garantir a robustez dos resultados, nesse trabalho realizamos as estimações por dois métodos: vizinho mais próximo e *kernel* – que são os dois mais utilizados pela literatura⁸⁵.

⁸⁵ O método pelo vizinho mais próximo consiste em, para cada unidade tratada, procurar a unidade de controle com o escore de propensão mais próximo (pode ser usado na forma “sem reposição”, o que significa que cada unidade de controle somente pode ser pareada uma única vez, ou “com reposição”, quando a mesma unidade de controle pode ser utilizada como par de mais de uma unidade tratada, modalidade que

Becker e Ichino (2002) destacam que a escolha por cada um dos métodos envolve um *trade-off* entre qualidade e quantidade de pareamentos. Nenhum deles pode ser apontado *a priori* como um método superior. Os autores aconselham considerá-los conjuntamente na análise, de forma a avaliar a robustez das estimativas. Trata-se da abordagem que adotaremos neste trabalho. Por outro lado, Caliendo & Kopeinig (2008) destacam a menor variância obtida por meio do método *kernel*, apontando-o como um método mais vantajoso.

3.4.2. Escolha das Variáveis - o Vetor de Características X

Como vimos, a estratégia de *matching* está fortemente amparada na hipótese de independência condicional. Em outras palavras, as variáveis de interesse devem ser independentes do tratamento condicional no escore de propensão. A escolha das variáveis de controle (X) deve, portanto, satisfazer essa condição. Somente variáveis que influenciam simultaneamente a decisão de participar e as variáveis de interesse devem ser incluídas.

Utilizaremos cinco grupos de variáveis como controles do pareamento, a saber: variáveis relativas a características acadêmicas da escola (ex. Ideb⁸⁶, taxa de aprovação, etc.), características físicas da escola (ex. existência de internet, biblioteca, laboratório de ciências, número de alunos, número de computadores, número de funcionários, etc.),

adotaremos nesse trabalho). Se denotarmos $C(i)$ o conjunto de controles pareados com a unidade de tratamento i , o pareamento pelo vizinho mais próximo irá calcular o pareamento segundo a seguinte regra:

$$C(i) = \min_j \|p_i - p_j\|$$

Já o pareamento Kernel consiste em comparar cada unidade tratada com uma média ponderada de todas as unidades de controle. O peso de cada controle nesse pareamento é inversamente proporcional à distância do seu escore de propensão e do escore de propensão de cada um dos tratados. No pareamento por escore de propensão, os pesos serão definidos por:

$$w(i, j) = \frac{K\left(\frac{p_j - p_i}{h}\right)}{\sum_{j=1}^N K\left(\frac{p_j - p_i}{h}\right)}$$

onde “K” é a função *kernel*, p_i representa o escore de propensão da unidade tratada, p_j os escores de propensão das unidades de controle e h é o parâmetro de largura da banda da função kernel.

⁸⁶ O Índice de Desenvolvimento da Educação Básica (Ideb) foi criado pelo Inep em 2007 e é resultado da combinação de dois outros indicadores: i) pontuação média dos estudantes em exames padronizados ao final de cada etapa do ensino fundamental (5º e 9º anos); e ii) taxa de aprovação dos estudantes em cada etapa de ensino.

participação em outros programas do MEC (ex. programa mais educação, programa de banda larga na escola, etc.), características do prefeito (alinhamento com o governo federal, estadual, se tem nível superior, etc.) e características do município (ex. PIB *per capita*, população, etc.). As variáveis de controle foram observadas no ano imediatamente anterior ao tratamento (2009).

Na Tabela A.1 do Apêndice, apresentamos e descrevemos brevemente as variáveis utilizadas no cálculo do escore de propensão.

3.4.3. Diferenças-em-Diferenças (DD)

A ideia fundamental do método de diferenças-em-diferenças é comparar amostras de tratados e não tratados antes e depois da intervenção. Sabemos que a simples comparação entre o desempenho acadêmico antes e depois do Prouca não permite a inferência quanto ao impacto causal, uma vez que outros fatores podem ter influenciado o desempenho nesse período (variável omitida). Além disso, não podemos simplesmente comparar o desempenho de alunos em escolas que aderiram ao Prouca e escolas que não aderiram, por conta do viés de autosseleção. Contudo, podemos combinar essas duas técnicas: comparar os resultados antes e depois de um grupo que se inscreveu e de um grupo que não se inscreveu no programa.

A estratégia consiste, portanto, em tirar “duas diferenças” (razão pela qual o método é chamado de diferença-em-diferença ou “dupla diferença”). Inicialmente, computamos a diferença nas médias da variável de interesse para cada um dos grupos nos dois períodos – essa é a chamada “primeira diferença”. Em seguida, de forma a controlar para a presença de fatores externos que podem ter afetado os dois grupos nesse período, tiramos uma “segunda diferença”: a diferença entre as diferenças calculadas para o grupo tratamento e para o grupo controle.

Assim, pelo método DD⁸⁷, construímos o contrafactual assumindo que o viés de seleção não varia com o tempo. Nesse caso, as mudanças na variável de interesse dos não-participantes no tratamento revelam as mudanças na variável de interesse dos participantes caso não houvesse tratamento. Em outras palavras, a variação do resultado no grupo de controle revela a variação contrafactual:

⁸⁷ Para a apresentação formal do método e sua aplicação, vide Card & Krueger (1994).

$$E(Y_1^T - Y_0^T | T_1 = 0) = E(Y_1^C - Y_0^C | T_1 = 0) \quad (4.3)$$

Note que no período 0 nenhum grupo foi tratado. Assim, $T_{0i}=0$ e $Y_{0i}=Y_{0i}^C$ para todo i .

O estimador de DD, que mensura os efeitos do tratamento nos tratados no período 1, será dado por:

$$DD = E(Y_1^T - Y_0^T | T_1 = 1) - E(Y_1^C - Y_0^C | T_1 = 0) = E(I_1 | T_1 = 1) \quad (4.4)$$

O primeiro elemento do lado esquerdo da equação representa a diferença na variável de interesse para o grupo dos tratados antes (Y_0^T) e depois (Y_1^T) do tratamento; o segundo elemento, representa a diferença na variável de interesse para o grupo dos controles antes (Y_0^C) e depois (Y_1^C) do tratamento. O estimador de diferença-em-diferença é nada mais do que a diferença entre esses dois elementos.

É importante notar que o método não exige que os dois grupos (tratamento e controle) sejam iguais – ou seja, os grupos não precisam ter as mesmas condições no momento pré-intervenção. É possível que o grupo de tratamento e de controle tenham não só características, como um desempenho de partida inferior/superior ao grupo de controle. A única exigência é que essas diferenças permaneçam constantes ao longo do tempo (na ausência do tratamento), ou seja, que se verifique a mesma tendência temporal, de forma a que as diferenças possam ser “controladas” ou “diferenciadas” (hipótese das tendências paralelas).

3.4.4. A Combinação dos Métodos: Escore de Propensão e Diferenças em Diferenças

Como vimos, tanto o método de diferenças em diferenças como o pareamento por escore de propensão possuem limitações. Uma forma de tentar minorá-las é por meio de uma combinação de ambos os métodos, de forma a aumentar a robustez do nosso contrafactual (grupo de controle).

O PSM possui uma limitação relevante: não é capaz de lidar com diferenças em características não observáveis. A técnica de diferença-em-diferença, por sua vez, não é capaz de controlar para mudanças nas características que afetem de forma desigual os grupos de tratamento e de controle. Uma combinação dos dois métodos permite, assim,

um desenho de avaliação superior: é capaz de lidar com características não observáveis que são constantes ao longo do tempo, ao mesmo tempo em que diminui as chances de existência de fatores que afetam desigualmente os dois grupos, ao criar, por meio do pareamento, um grupo de controle com características semelhantes ao grupo de tratamento. Em outras palavras, o método de PSM com DD é capaz de lidar, simultaneamente, com o viés de seleção nas características observáveis e não observáveis fixas no tempo, relaxando a hipótese de seleção em observáveis do PSM.

Como destacam Hirano, Imbens, & Ridder (2003), o uso conjunto do DD com o PSM possibilita a estimação por meio de uma regressão ponderada pelo escore de propensão de cada observação. A ponderação busca corrigir distorções do modelo DD, produzindo um estimador eficiente. Assim, controlamos para variáveis observáveis utilizando o PSM e para variáveis não observáveis invariantes no tempo utilizando a técnica de diferenças-em-diferenças.

Nota-se que a combinação ainda possui limitações, uma vez que não é capaz de controlar para características não observadas que afetem desigualmente o grupo de tratamento e de controle. Contudo, o pareamento diminui a probabilidade de sua existência, uma vez que criamos um grupo de controle formado por unidades com características bastante similares àquelas encontradas nas unidades do grupo de tratamento.

Neste trabalho, a combinação das duas técnicas foi feita em dois passos: i) inicialmente realizamos o pareamento com base em características de linha de base (antes do tratamento); ii) aplicamos o método de diferenças-em-diferenças somente nas unidades pareadas. Dessa forma, para cada um dos grupos de interesse foi definido um ou mais grupos de controle formado por escolas com características semelhantes que não aderiram ao Prouca (os grupos de controle variaram conforme o método de pareamento adotado).

3.5 Base de Dados

3.5.1. Fontes Utilizadas

Para construção das estimativas desse trabalho, utilizamos quatro bases de dados principais: Prova Brasil e o Censo Escolar

A Prova Brasil consiste em uma avaliação censitária da rede pública realizada pelo Inep. As provas são aplicadas para estudantes do 5º e do 9º ano do Ensino

Fundamental a cada dois anos. Utilizamos os dados da Prova Brasil para mensuração de quatro variáveis de interesse: desempenho acadêmico dos alunos do 5º e 9º anos em Português e Matemática e percentual de professores que utilizam o computador e a internet para fins pedagógicos. Os dados foram analisados para o período entre 2007 e 2015. Medimos, ainda, os resultados do Prouca por meio de outras duas variáveis: o Ideb⁸⁸ e as taxas de abandono. Os dados foram obtidos diretamente no site do Instituto.

Utilizamos, ademais, os dados do Censo Escolar, que é o principal instrumento de coleta de informações da educação básica e abrange diferentes etapas de ensino. Trata-se de um levantamento de dados estatísticos educacionais realizado, anualmente, em âmbito nacional, sob coordenação do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - Inep. Participam do Censo todas as escolas públicas e privadas do país. Neste trabalho, analisamos os dados do Censo Escolar entre 2007 e 2016, com enfoque nas informações relativas às escolas e às matrículas de estudantes da rede pública. Excluimos os dados relativos às escolas privadas, escolas que não estavam em funcionamento em determinado ano e escolas destinadas a portadores de deficiência. Além disso, excluimos os dados relativos a escolas federais, por se tratar de um conjunto de escolas públicas com características bastante peculiares. Os dados do Censo Escolar foram utilizados, principalmente, para a mensuração das variáveis de controle.

Complementarmente, utilizamos dados fornecidos pela Secretaria de Educação Básica do Ministério da Educação. Essa base de dados é focada em escolas participantes do Prouca em 2010 e traz informações relativas a: i) escolas que participaram do programa; e ii) número de *laptops* entregues por escola. Além disso, obtivemos junto a Secretaria bases relativas a participação de escolas em quatro outros programas do Ministério: Programa Nacional de Tecnologia Educacional (ProInfo), Programa Nacional de Banda Larga (PNBL), Programa Governo Eletrônico - Serviço de Atendimento ao Cidadão (Gesac) e Programa Mais Educação (PME). As quatro bases foram utilizadas para mensuração de variáveis de controle (se as escolas participaram ou participavam desses programas antes do Prouca).

⁸⁸ O Índice de Desenvolvimento da Educação Básica (Ideb) foi criado pelo Inep em 2007 e é resultado da combinação de dois outros indicadores: i) pontuação média dos estudantes em exames padronizados ao final de cada etapa do ensino fundamental (5º e 9º anos); e ii) taxa de aprovação dos estudantes em cada etapa de ensino.

Finalmente, lançamos mão de dados do Instituto Brasileiro de Geografia e Estatística – IBGE para os dados de PIB Municipal, população do município e para informações a respeito do prefeito, e do Programa das Nações Unidas para o Desenvolvimento – PNUD para os dados relativos ao Índice de Desenvolvimento Humano Municipal. Esses três conjuntos de dados também foram utilizados para a criação de variáveis de controle para realização do pareamento.

3.5.2. Estatísticas Descritivas

Inicialmente, apresentaremos estatísticas descritivas para o grupo tratamento e o grupo controle considerando a amostra de Anos Iniciais⁸⁹ e Anos Finais⁹⁰ do Ensino Fundamental. As estatísticas descritivas para a amostra relativa aos Anos Iniciais podem ser observadas na Tabela 19 e para a amostra relativa aos Anos Finais na Tabela A.2. do Apêndice.

Conforme dados obtidos junto ao Ministério da Educação, 367 escolas receberam computadores do Prouca em 2010. Dessas, 346 foram identificadas no Censo Escolar 2009. Para a amostra dos Anos Iniciais, excluimos as escolas que não possuíam alunos matriculados no 1º ao 5º ano do Ensino Fundamental regular e escolas que não possuíam dados relativos ao SAEB nos anos analisados. Dessa forma, nossa base para os anos iniciais é composta por 26.180 escolas no grupo de controle e 149 escolas no grupo tratamento. Para os anos finais, foi adotado procedimento semelhante, o que resultou em um grupo de controle formado por 21.744 escolas e um grupo tratamento constituído por 95 escolas.

Como se nota, há diferenças significativas nas médias e nas distribuições de diversas variáveis entre o grupo controle e o grupo tratamento antes da implementação do Prouca, o que indica que as escolas dos dois grupos são diferentes entre si e demonstra as vantagens de se proceder o pareamento.

⁸⁹ Os anos iniciais correspondem ao período entre o 1º e o 5º ano. A avaliação dos anos iniciais é feita por meio da aplicação da Prova Brasil aos alunos do 5º ano das escolas.

⁹⁰ Os anos finais correspondem ao período entre o 6º e o 9º ano. A avaliação dos anos finais é feita por meio da aplicação da Prova Brasil aos alunos do 9º ano das escolas.

Tabela 19 - Estatísticas Descritivas e testes de médias – Anos Iniciais

	Escolas Tratadas = 149		Escolas Controle = 26.180		Teste de Igualdade de Médias	
	Média	Desvio padrão	Média	Desvio padrão	Diferença	p-valor
Prova Brasil						
Nota_5F_Port	179.65	19.34	182.31	20.63	-2.67	0.12
Nota_5F_Mat	199.33	21.92	202.66	24.60	-3.33	0.10
Ideb_Anos_Iniciais	4.44	0.93	4.59	1.05	-0.15	0.09
Abandono_Anos Iniciais	1.85	3.17	1.91	3.15	-0.06	0.82
Aprovação_Anos_Iniciais	88.03	8.81	88.68	9.02	-0.65	0.38
TDI_Anos_Iniciais	19.47	11.16	18.27	12.45	1.20	0.24
Censo Escolar 2009						
id_dependencia_adm	0.30	0.46	0.24	0.43	0.05	0.13
id_internet	0.76	0.43	0.74	0.44	0.02	0.53
id_banda_larga	0.55	0.50	0.58	0.49	-0.03	0.46
id_biblioteca	0.66	0.48	0.57	0.50	0.09	0.02
id_quadra_esportes	0.51	0.50	0.55	0.50	-0.04	0.29
num_salas_existentes	10.18	4.94	11.05	5.52	-0.86	0.06
num_computadores2009	15.74	39.86	10.82	10.57	4.92	0.00
computador_por_aluno2009	0.03	0.06	0.02	0.02	0.01	0.00
id_laboratorio_ciencias	0.09	0.28	0.10	0.30	-0.01	0.60
id_laboratorio_informatica	0.66	0.48	0.58	0.49	0.08	0.04
id_sala_diretoria	0.89	0.31	0.93	0.25	-0.04	0.07
id_sala_professor	0.90	0.30	0.84	0.36	0.06	0.06
id_equip_impresora	0.93	0.25	0.93	0.26	0.00	0.87
num_funcionarios	47.21	22.75	48.10	26.78	-0.89	0.70

numero_de_alunos2009	469.87	268.62	576.61	352.95	-106.74	0.00
participantes_5F_2009	51.40	25.75	69.70	44.79	-18.30	0.00
tem_ef_finais	0.68	0.47	0.53	0.50	0.15	0.00
tem_em	0.05	0.23	0.09	0.29	-0.04	0.11
tem_ed_infantil	0.42	0.49	0.39	0.49	0.03	0.51
<hr/>						
Dados_MEC						
proinfo	0.72	0.45	0.68	0.47	0.04	0.27
gesac	0.09	0.28	0.04	0.19	0.05	0.00
pnbl	0.83	0.37	0.79	0.41	0.05	0.18
pme	0.09	0.29	0.10	0.30	-0.01	0.71
<hr/>						
IBGE_Prefeitos						
partido_pt_2009	0.15	0.36	0.16	0.36	0.00	0.92
partido_oposicao_2009	0.17	0.37	0.25	0.43	-0.08	0.02
prefeito_ensino_superior	0.57	0.50	0.67	0.47	-0.10	0.01
sexo_prefeito	0.90	0.30	0.90	0.30	0.00	0.93
alinhamento_governador	0.23	0.43	0.21	0.41	0.02	0.54
<hr/>						
IBGE_Municipios						
pib_pcap2009	14,114	9,847	15,757	14,488	-1,642	0.17
populacao2009	353,748	1,065,433	654,962	1,802,077	-301,213	0.04
idhm	0.70	0.08	0.71	0.08	-0.01	0.09

Fonte: Elaboração própria a partir dos dados pesquisados.

3.6 Análise dos Resultados

3.6.1. Balanceamento do Escore de Propensão

As variáveis de controle para estimação do escore de propensão foram selecionadas a partir de análises da literatura e de fatores que poderiam ter influenciado a escolha de uma escola para participação no programa. Foram testadas várias especificações a partir das variáveis descritas na seção 3.4. O modelo completo mostrou-se o mais apropriado quando comparado com modelos que contemplavam a inserção de um ou mais blocos de variáveis.

As Tabelas A.3 e A.4 do Apêndice trazem os resultados do pareamento realizado por meio do vizinho mais próximo para a amostra relativa aos Anos Iniciais e aos Anos Finais. Como se nota, o pareamento foi realizado com êxito. Tanto para os Anos Iniciais quanto para os Anos Finais, não se observa diferença significativa entre as escolas do grupo tratamento e as escolas pareadas do grupo controle em nenhuma das variáveis analisadas.

Um indício importante do sucesso do pareamento pode ser obtido pela comparação entre as tendências pré-tratamento dos grupos tratamento e controle, que pode ser observada para as principais variáveis na Figura 20. Como se nota, tanto para os anos iniciais quanto para os anos finais, os dois grupos de escola seguiram trajetórias não apenas paralelas, mas virtualmente iguais no período pré-Prouca. Esses gráficos serão analisados com maior detalhe nas seções seguintes.

3.6.2. Impacto no número de computadores e de computadores por aluno

Inicialmente, avaliamos o impacto do Prouca no número de computadores por escola e por aluno medidos a partir dos dados do Censo Escolar. A cada ano, as escolas devem preencher o questionário do Censo referente ao ano anterior. Uma das perguntas a serem respondidas é relativa a “quantidade de computadores na escola”. A análise do impacto do programa nessa variável do Censo é importante com vistas a assegurar que o tratamento foi, de fato, implementado.

A Tabela 20 e a Tabela 21 trazem os resultados dessas estimativas. As variáveis na forma “D.2009__(ano)” referem-se às diferenças na variação do número de

computadores (e do número de computadores por aluno) entre 2009 e o ano de referência para os dois grupos de escolas (tratamento e controle).

A título ilustrativo, no caso da variável “computadores por escola”, o coeficiente de $D.2009_2011$ representa a diferença na variação média do número de computadores por escola entre o grupo tratamento e o grupo controle entre 2009 e 2011 (ou seja, é o estimador de DD após o pareamento). Adotaremos esta notação ao longo de todo o trabalho.

As tabelas reportam, assim, os resultados do impacto do Prouca no número de computadores e no número de computadores por aluno de 2010 a 2016 utilizando três métodos: diferenças-em-diferenças (DD) sem pareamento, diferenças-em-diferenças com pareamento *kernel* ($PSM + DD - kernel$) e com pareamento pelo método do vizinho mais próximo – *nearest neighbour* ($PSM + DD - NN$).

Como se nota, os resultados do programa aparecem já em 2010, com uma diferença de crescimento de 22 -24 computadores por escola em relação a 2009 nos anos iniciais e de 35-45 nos anos finais. Já a partir de 2011, essa diferença torna-se ainda mais significativa, de 161-165 computadores nos anos iniciais e de 97-111 computadores nos anos finais. Os resultados indicam, assim, que os computadores foram realmente inseridos na rotina escolar da maioria das escolas somente a partir de 2011.

Chama a atenção o fato de que, a partir de 2014, essa diferença começa a cair significativamente, ao ponto de, em 2015, para os anos finais, e em 2016, para os anos iniciais, não se verificar qualquer diferença significativa entre o número de computadores por escola no grupo tratado e no grupo controle. Essa evolução fica mais explícita quando analisamos a Figura 16 e a Figura 17 . Como se nota, após um período de dois anos (2011 e 2012) nos quais a diferença no número de computadores atinge seu auge, observa-se uma forte redução. Os dados indicam, assim, a rápida inutilização (seja por furto ou roubo, seja por problemas técnicos) dos computadores Prouca – que tem início após apenas 2 anos de utilização.

Tabela 20 – Estimativas de Impacto do Prouca no número de computadores por escola e por aluno – Anos Iniciais

Variável	Anos Iniciais			Anos Iniciais		
	Computadores por Escola			Computador por Aluno		
	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)
D.2009_2010	25.37*** (1.48)	22.05** (9.08)	24.16*** (9.71)	0.05*** (0)	0.04** (0.02)	0.05*** (0.02)
D.2009_2011	164.36*** (2.02)	161.46*** (18.89)	165.1*** (18.95)	0.35*** (0)	0.34*** (0.04)	0.35*** (0.04)
D.2009_2012	163.35*** (2.3)	160.79*** (18.67)	164.18*** (18.73)	0.33*** (0)	0.33*** (0.04)	0.34*** (0.04)
D.2009_2013	119.36*** (2.7)	122.95*** (17.54)	119.88*** (17.85)	0.25*** (0.01)	0.25*** (0.04)	0.25*** (0.04)
D.2009_2014	51.53*** (1.77)	55.4*** (12.74)	54.93*** (13.14)	0.11*** (0)	0.12*** (0.03)	0.12*** (0.03)
D.2009_2015	9.76*** (2.3)	13.17** (6.15)	12.7** (6.89)	0.02*** (0)	0.03** (0.01)	0.03 (0.02)
D.2009_2016	1.30 (1.87)	5.23 (4.14)	2.47 (5.33)	0.00 (0)	0.01 (0.01)	0.01 (0.01)
Total de Observações:	26.329	26.329	26.329	26.329	26.329	26.329

Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%.

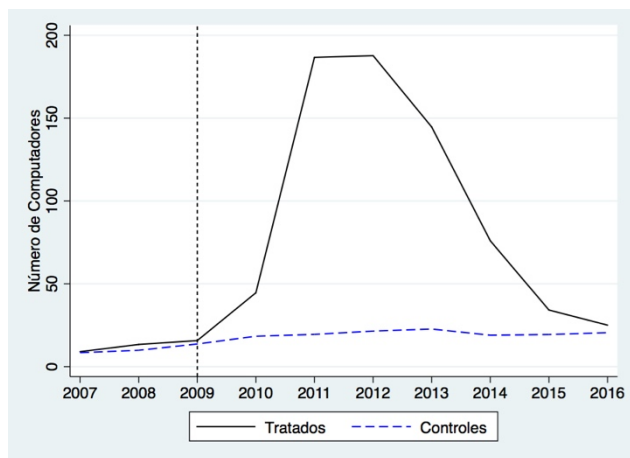
Tabela 21 – Estimativas de Impacto do Prouca no número de computadores por escola e por aluno – Anos Finais

Variável	Anos Finais			Anos Finais		
	Computadores por Escola			Computador por Aluno		
	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)
D.2009_2010	39.83*** (2.15)	34.99** (18.82)	45.77** (19.93)	0.05*** (0)	0.05** (0.02)	0.06** (0.02)
D.2009_2011	104.78*** (2.43)	97.1*** (23.85)	110.52*** (24.57)	0.19*** (0)	0.18*** (0.04)	0.20*** (0.04)
D.2009_2012	128.84*** (2.93)	122.00*** (25)	135.4*** (25.49)	0.23*** (0)	0.23*** (0.04)	0.24*** (0.04)
D.2009_2013	73.6*** (2.53)	76.05*** (22.17)	79.6*** (23.1)	0.14*** (0)	0.14*** (0.04)	0.14*** (0.04)
D.2009_2014	12.73*** (1.99)	19.68 (10.51)	19.57 (12.6)	0.03*** (0)	0.04 (0.02)	0.04 (0.02)
D.2009_2015	-0.93 (3.05)	5.76 (6.3)	4.97 (9.44)	0.01** (0.005)	0.02 (0.02)	0.02 (0.02)
D.2009_2016	-5.8** (2.45)	0.80 (4.21)	1.17 (8.32)	0.00 (0)	0.01 (0.01)	0.00 (0.02)
Total de Observações:	21.839	21.839	21.839	21.839	21.839	21.839

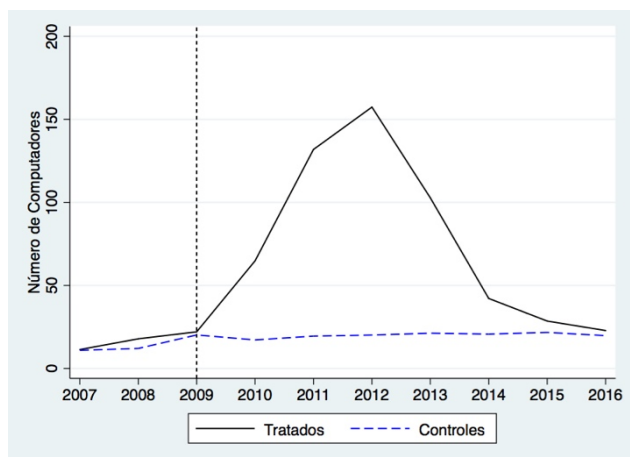
Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%.

Figura 16 - Impacto do Prouca no Número de Computadores por Escola

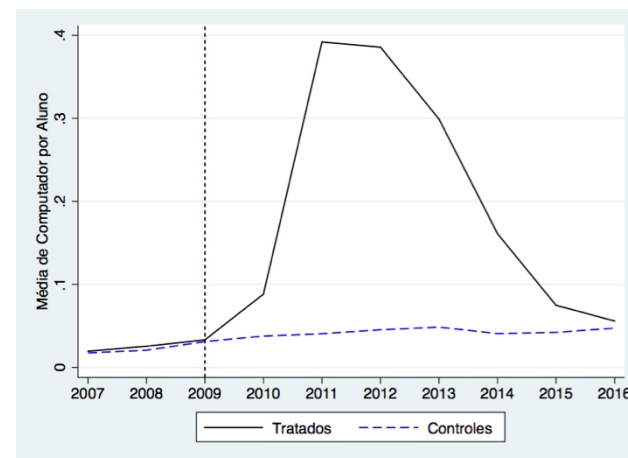


(a) Anos Iniciais

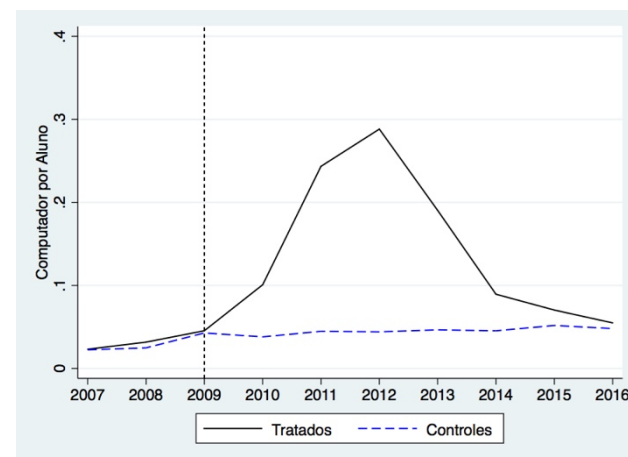


(b) Anos Finais

Figura 17 - Impacto do Prouca no Número de Computadores por Aluno



(a) Anos Iniciais



(b) Anos Finais

Os dados relativos ao número de computadores por aluno apresentam quadro semelhante. Nesse caso, chama a atenção o fato de a razão computador por aluno reportada no Censo Escolar não ter ultrapassado, em nenhum momento, 0,39 (anos iniciais) e 0,29 (anos finais). Uma análise da base de dados mostra que mais de 50% das escolas tratadas não reportaram aumento expressivo na quantidade de computadores na escola no Censo Escolar em nenhum dos anos entre 2010 e 2016.

Uma hipótese é que esses computadores, apesar de enviados pelo MEC, nunca tenham chegado nas escolas ou nunca tenham sido, de fato, utilizados. Uma explicação alternativa – que julgamos mais provável⁹¹ – é que os *laptops* Prouca não fossem compreendidos por aqueles que preenchem o Censo Escolar como “computadores da Escola”, seja pelo fato de estarem vinculados aos alunos seja por uma visão mais tradicional do que seria um computador (*desktop*).

3.6.3. Impacto na utilização de computadores e internet para fins pedagógicos

Avaliamos, em seguida, o impacto do Prouca no percentual do corpo docente da escola que declarou utilizar computador e internet para fins pedagógicos. Os dados foram retirados do questionário da Prova Brasil. Em virtude de falhas no preenchimento do questionário, para essa variável nossos grupos de tratamento são menores: 116 para os anos iniciais e 73 para os anos finais. O número de escolas no grupo controle também diminuiu, mas permanece bastante expressivo (21.221 escolas nos anos iniciais e 17.659 nos anos finais). Nesse caso, o pareamento kernel mostrou-se superior ao pareamento por vizinho mais próximo.

Como se nota na Tabela 22, o programa teve impacto positivo e estatisticamente significativo nos anos imediatamente posteriores a sua implementação para os anos iniciais. De 2009 para 2011, verificamos um aumento de 19-22 pontos percentuais no total de professores que declararam utilizar o computador para fins pedagógicos. Movimento semelhante ocorre com a utilização da internet. Nesse caso, verificou-se um aumento de 16-18 pontos percentuais.

⁹¹ A razão de julgarmos mais provável é que, em nenhum dos estudos de caso do Prouca analisados, encontramos exemplos em que as escolas não tenham recebido os *laptops* previstos.

Nos dois casos, contudo, há indícios de que esse impacto foi de curto prazo, tendo sido revertido nos anos seguintes. Já em 2015, não se verifica qualquer diferença estatisticamente significativa no uso do computador e da internet para fins pedagógicos entre escolas tratadas e escolas do grupo de controle.

Já para os anos finais, o quadro é ligeiramente distinto. De 2009 a 2011, verificamos uma diferença de evolução de 11-14 pontos percentuais entre escolas tratadas e controle nos professores que declararam utilizar o computador para fins pedagógicos. Esse efeito foi também de curto prazo: já em 2013 não se verifica qualquer diferença na evolução entre os dois grupos de escolas. No caso da internet, não se verificou qualquer impacto significativo no curto prazo em nenhuma das estimações. Em 2015, observou-se um impacto negativo do programa nas estimações sem pareamento e com pareamento *kernel*, mas não significativo na estimacão por vizinho mais próximo

Os efeitos do programa na utilização pelos docentes de computadores e da internet para fins pedagógicos podem ser observados, ainda, na Figura 18 e na Figura 19.

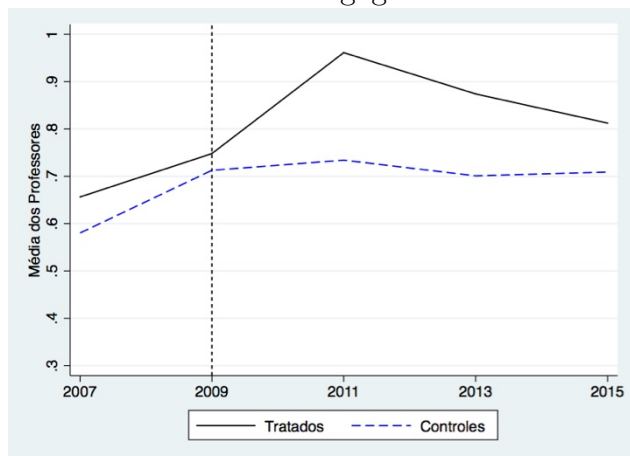
Tabela 22 – Estimativas de Impacto do Prouca no uso de computadores e internet para fins pedagógicos – Anos Iniciais

Variável	Anos Iniciais					
	Usa Computador			Usa Internet		
	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)
D.2009_2011	0.19*** (0.04)	0.19*** (0.04)	0.22*** (0.06)	0.16*** (0.05)	0.16*** (0.04)	0.18*** (0.06)
D.2009_2013	0.14*** (0.05)	0.14*** (0.04)	0.10 (0.06)	0.09 (0.05)	0.09** (0.04)	0.06 (0.06)
D.2009_2015	0.07 (0.05)	0.07 (0.04)	0.10 (0.06)	-0.03 (0.05)	-0.03 (0.05)	-0.04 (0.07)
Total de Observações:	21.337	21.337	21.337	21.337	21.337	21.337
Variável	Anos Finais					
	Usa Computador			Usa Internet		
	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)
D.2009_2011	0.11** (0.05)	0.11*** (0.04)	0.14** (0.07)	0.05 (0.05)	0.05 (0.05)	0.08 (0.07)
D.2009_2013	-0.05 (0.05)	-0.05 (0.05)	-0.01 (0.08)	-0.02 (0.05)	-0.02 (0.05)	0.03 (0.07)
D.2009_2015	-0.06 (0.05)	-0.06 (0.05)	-0.02 (0.08)	-0.14*** (0.05)	-0.13*** (0.05)	-0.05 (0.08)
Total de Observações:	17.732	17.732	17.732	17.732	17.732	17.732

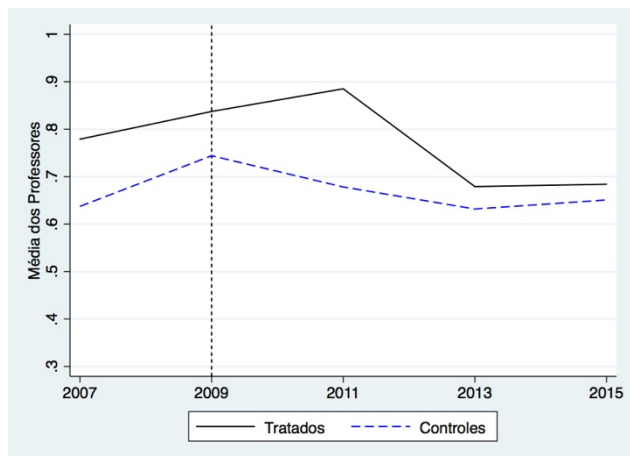
Fonte: Elaboração própria a partir dos dados pesquisados.

*** Significativo a 1%. ** Significativo a 5%.

Figura 18 - Impacto do Prouca no Uso do Computador Para Fins Pedagógicos

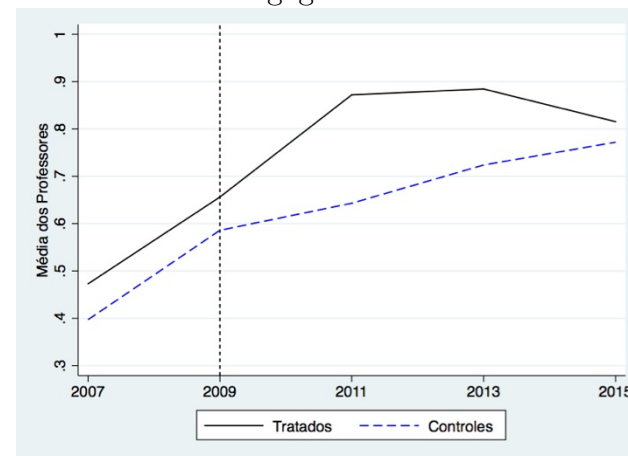


(a) Anos Iniciais

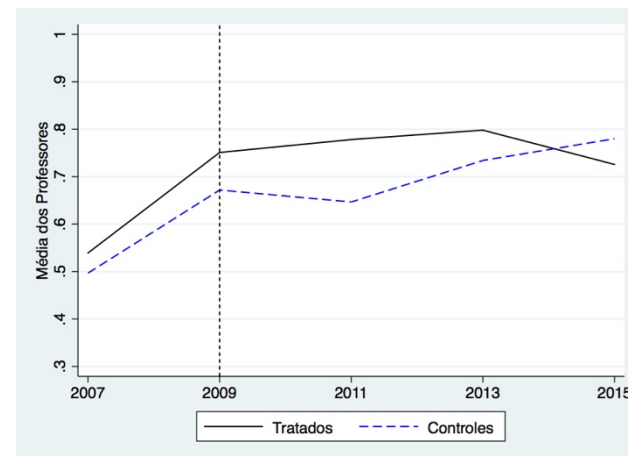


(b) Anos Finais

Figura 19 - Impacto do Prouca no Uso da Internet Para Fins Pedagógicos



(a) Anos Iniciais



(b) Anos Finais

3.6.4. Impacto no desempenho acadêmico dos estudantes e nas taxas de abandono

Nesta seção, iremos analisar o impacto do Prouca no desempenho acadêmico dos estudantes, medido por meio das provas de matemática e português do SAEB. Avaliaremos, ainda, o impacto no Ideb das escolas e nas taxas de abandono. Como vimos na seção 3.2, a literatura sobre o impacto dos programas 1:1 no desempenho acadêmico dos estudantes pode ser considerada mista, com uma prevalência de efeitos nulos, principalmente em países em desenvolvimento. Esse é justamente o caso do Prouca.

A Tabela 23 traz as estimativas de impacto do Prouca no desempenho acadêmico dos estudantes. No tocante ao desempenho nas provas de português, não foi constatada qualquer evidência de impacto do programa nos anos iniciais. Já nos anos finais, a estimação por meio de diferenças-em-diferenças sem pareamento encontrou evidências de impacto negativo do Prouca em 2011 e 2013, a nível de significância de 5%. Em 2013, também a estimação por kernel encontrou evidências de impacto negativo. Contudo, essas diferenças não se mostraram relevantes na estimativa com o pareamento realizado pelo vizinho mais próximo, que consideramos mais robusta. Assim, nossas estimativas indicam que o programa teve impacto nulo nas notas de português nos anos iniciais e nos anos finais. No caso do desempenho em matemática, não encontramos evidência de impacto significativo do Programa nem nos anos iniciais nem nos anos finais em nenhuma das especificações analisadas, tanto a curto quanto a médio e longo prazo.

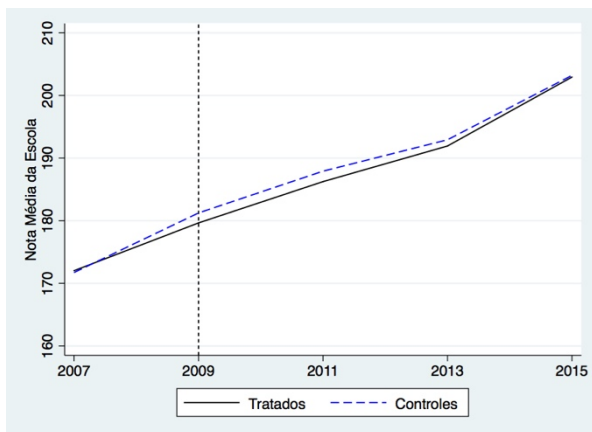
Quando avaliamos o impacto no Ideb da escola, o quadro é, naturalmente, parecido. O Prouca não apresentou qualquer impacto nos anos iniciais, em nenhuma das estimações. Já nos anos finais, encontramos evidência de um pequeno impacto negativo em 2013 a 5% na estimação sem pareamento. Contudo, as estimativas mais robustas, com pareamento, apontam para a ausência de impacto significativo do tratamento.

Tabela 23 – Estimativas de Impacto do Prouca no desempenho acadêmico e no abandono – Anos Iniciais e Finais

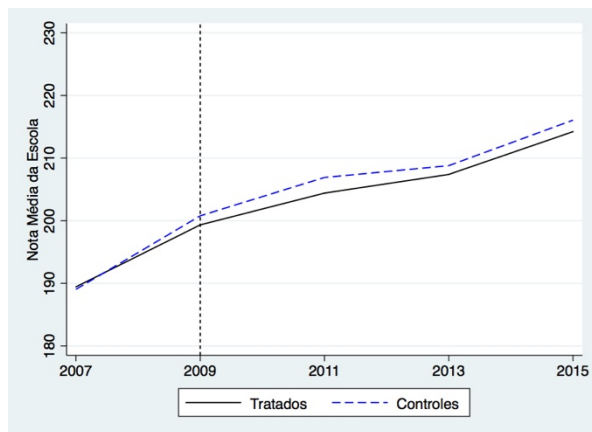
Variável	Anos Iniciais											
	Português			Matemática			IDEB			Abandono		
	DD (sem (pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem (pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem (pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem (pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)
D.2009_2011	0.19 (1.1)	0.31 (1.2)	-0.07 (1.67)	-0.24 (1.29)	-0.22 (1.34)	-1.05 (1.92)	0.03 (0.05)	0.07 (0.06)	-0.02 (0.07)	-0.14 (0.2)	-0.15 (0.22)	-0.06 (0.3)
D.2009_2013	1.45 (1.3)	1.50 (1.44)	0.60 (1.89)	1.53 (1.49)	1.58 (1.69)	0.05 (2.29)	0.07 (0.05)	0.07 (0.06)	-0.02 (0.08)	-0.14 (0.21)	-0.15 (0.22)	-0.01 (0.33)
D.2009_2015	0.74 (1.34)	0.99 (1.32)	1.33 (2.04)	0.33 (1.53)	0.44 (1.53)	-0.39 (2.38)	0.04 (0.06)	0.04 (0.06)	0.00 (0.09)	-0.16 (0.23)	-0.18 (0.23)	-0.14 (0.36)
Total de Observações:	26.329	26.329	26.329	26.329	26.329	26.329	26.329	26.329	26.329	26.329	26.329	26.329
Variável	Anos Finais											
	Português			Matemática			IDEB			Abandono		
	DD (sem (pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem (pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem (pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)	DD (sem (pareamento)	PSM + DD (<i>kernel</i>)	PSM + DD (NN)
D.2009_2011	-3.09** (1.44)	-2.94 (1.66)	-0.49 (2.23)	-2.56 (1.39)	-2.67 (1.72)	-0.36 (2.1)	-0.07 (0.06)	-0.06 (0.07)	0.02 (0.09)	-0.75 (0.45)	-0.73 (0.5)	-0.73 (0.62)
D.2009_2013	-3.81** (1.62)	-3.48** (1.61)	1.16 (2.47)	-2.58 (1.57)	-2.26 (1.73)	1.42 (2.48)	-0.14** (0.07)	-0.12 (0.07)	-0.07 (0.11)	-0.63 (0.5)	-0.64 (0.6)	-0.07 (0.76)
D.2009_2015	-2.95 (1.76)	-2.20 (1.73)	-1.59 (2.49)	-1.16 (1.64)	-0.64 (1.75)	0.34 (2.47)	-0.09 (0.07)	-0.07 (0.08)	-0.07 (0.11)	-0.19 (0.52)	-0.22 (0.59)	0.23 (0.75)
Total de Observações:	21.839	21.839	21.839	21.839	21.839	21.839	21.839	21.839	21.839	21.839	21.839	21.839

Fonte: Elaboração própria a partir dos dados pesquisados. *** Significativo a 1%. ** Significativo a 5%.

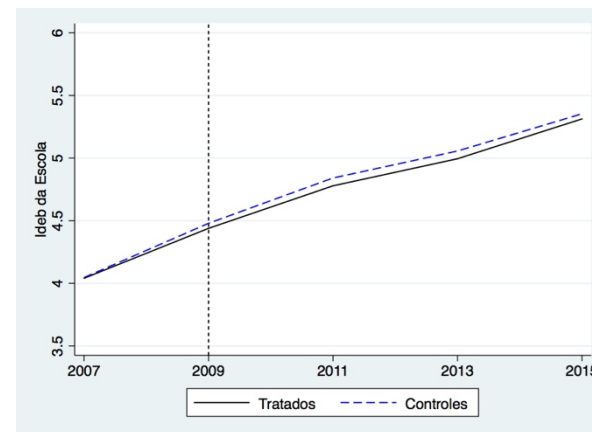
Figura 20 - Impactos do Procura no Desempenho Acadêmico



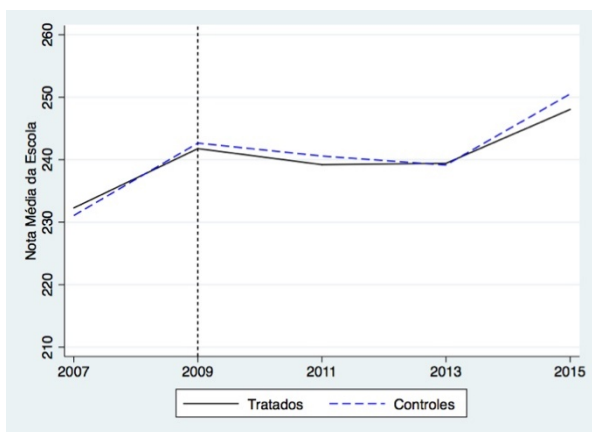
(a) Nota de Português - Anos Iniciais



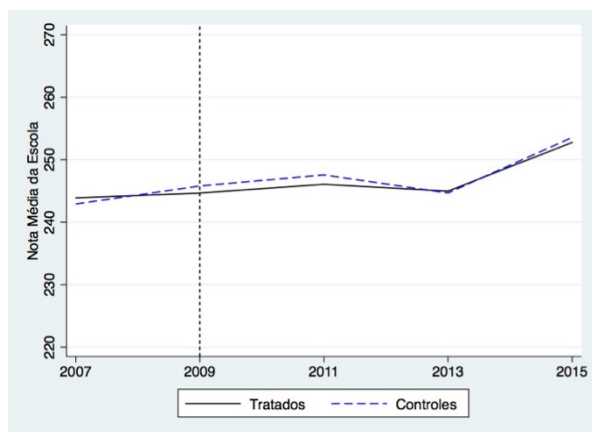
(b) Nota de Matemática - Anos Iniciais



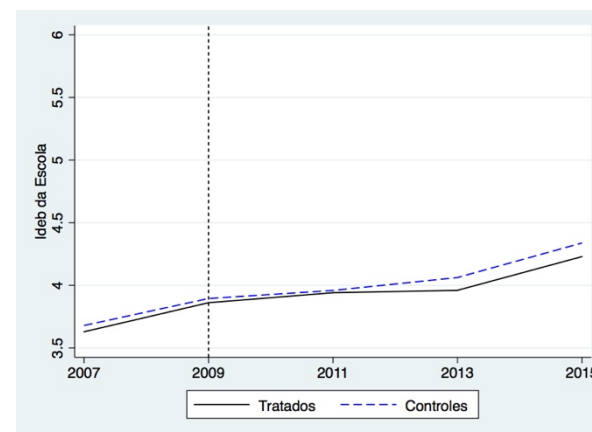
(c) IDEB - Anos Iniciais



(d) Nota de Português - Anos Finais



(e) Nota de Matemática - Anos Finais



(f) IDEB - Anos Finais

Como se nota, o cenário é bastante negativo para o programa. O Prouca não impactou, de forma significativa, o desempenho acadêmico dos estudantes em português e em matemática, assim como não teve qualquer efeito no Ideb das escolas contempladas com *laptops* do Programa. A ausência de impacto do programa pode ser observada com clareza na Figura 20, que apresenta a evolução das variáveis de interesse para o grupo tratamento e para as escolas pareadas do grupo controle entre 2007 e 2015. Fica evidente que o programa não alterou as trajetórias do grupo tratado em nenhuma das variáveis relativas ao desempenho acadêmico e em nenhum dos anos avaliados.

Uma hipótese levantada por alguns estudos (Trimmel & Bachmann (2004); Kozma et al (2004) é que a inserção da tecnologia no ambiente escolar aumentaria a “motivação/empenho” dos alunos. Avaliamos essa assertiva analisando o impacto do Prouca nas taxas de abandono da escola. A hipótese é que a introdução dos *laptops* no ambiente escolar aumentaria a motivação dos alunos e, conseqüentemente, diminuiria o abandono. Contudo, essa hipótese não é corroborada pelas estimações. Como se nota nas últimas colunas da Tabela 23, o programa não impactou as taxas de abandono em nenhum dos anos e em nenhuma das especificações avaliadas, tanto nos anos iniciais quanto nos anos finais.

Os resultados estimados nesta seção são, portanto, coerentes com diversos outros artigos revisados na seção 3.2, tais como Cristia et al. (2012) e De Melo et al. (2014b), no sentido de demonstrar a ausência de impacto significativo de programas de larga escala 1:1 no desempenho acadêmico de estudantes em países em desenvolvimento.

3.6.5. Hipóteses Paralelas e Teste de Igualdade de Médias antes do Tratamento

De forma a verificar a robustez dos resultados estimados, realizamos um teste de igualdade de médias entre as escolas tratadas e as escolas pareadas do grupo controle antes do tratamento. Nosso objetivo foi assegurar que antes do Prouca as variáveis de interesse de ambos os grupos seguissem as mesmas trajetórias.

Deve-se notar que a hipótese das trajetórias paralelas não exige a ausência de diferença significativa entre as variáveis de interesse antes do tratamento, mas, simplesmente, como o próprio nome permite inferir, que elas sigam trajetórias paralelas. O teste de igualdade de médias pode ser considerado, assim, uma exigência adicional, que contribui para assegurar a robustez dos resultados estimados.

Como se nota pela análise da Tabela 24, o grupo tratamento e controle pareado seguiram trajetórias não apenas paralelas, mas estatisticamente iguais entre 2007 e 2009, tanto para o grupo tratamento como para o grupo controle, o que sugere uma estimativa dos impactos do Prouca sem viés. Essa conclusão é corroborada pela análise da Figura 20, nos quais a trajetória do grupo de tratamento e controle entre 2007 e 2009 apresentam-se completamente sobrepostas para os anos iniciais e muito próximas - a ponto de serem estatisticamente insignificantes - para os anos finais. A análise da evolução das variáveis de interesse antes do tratamento corrobora, assim, a qualidade do pareamento e, conseqüentemente, contribui para a robustez dos resultados estimados.

Tabela 24 – Teste de Igualdade de Médias antes do Prouca

	Anos Iniciais			
	Escolas Tratadas	Escolas Controle	Teste de Igualdade de Médias	
	Média	Média	Diferença	p-valor
2007				
Nota_5F_Port	172.02	171.72	0.30	0.89
Nota_5F_Mat	189.43	189.04	0.39	0.87
Ideb_Anos_Iniciais	4.04	4.05	-0.01	0.96
Abandono_Iniciais	3.13	4.05	-0.92	0.94
2009				
Nota_5F_Port	179.65	181.24	-1.59	0.48
Nota_5F_Mat	199.33	200.78	-1.45	0.58
Ideb_Anos_Iniciais	4.44	4.48	-0.04	0.72
Abandono_Iniciais	1.85	2.10	-0.25	0.51
	Anos Finais			
	Escolas Tratadas	Escolas Controle	Teste de Igualdade de Médias	
	Média	Média	Diferença	p-valor
2007				
Nota_9F_Port	232.28	231.07	1.21	0.67
Nota_9F_Mat	243.88	242.91	0.97	0.75
Ideb_Anos_Finais	3.63	3.68	-0.05	0.69
Abandono_Finais	5.19	6.56	0.25	-1.36
2009				
Nota_9F_Port	241.78	242.66	-0.88	0.75
Nota_9F_Mat	244.67	245.80	-1.13	0.71
Ideb_Anos_Finais	3.86	3.89	-0.03	0.79
Abandono_Finais	5.17	4.99	0.52	0.18

Fonte: Elaboração própria a partir dos dados pesquisados. *** Significativo a 1%. ** Significativo a 5%.

3.6.6. Resultados Heterogêneos

Foram realizadas, ainda, estimativas de efeitos heterogêneos do Prouca conforme as características do município (baixo ou alto PIB *per capita*). Em síntese, o quadro geral de avaliação do Programa não se altera com análise de heterogeneidade. A ausência de impacto do Prouca independe do PIB *per capita* dos municípios das escolas contempladas.

Realizamos, ainda, diversos outros testes de robustez: i) avaliamos o programa mantendo no grupo de controle somente escolas de municípios em que pelo menos uma escola recebeu o Prouca; ii) retiramos do grupo de tratamento os municípios do Uca-Total; iii) retiramos da amostra as escolas localizadas em capitais dos estados; iv) retiramos da amostra as escolas localizadas em áreas rurais; v) utilizamos como grupo de controle as escolas que receberam o Proinfo, mas não receberam Prouca; vi) avaliamos o impacto somente em escolas que tinham sido contempladas, até 2009, com o programa banda larga na escola; e vii) mantivemos no grupo tratamento apenas escolas que reportaram aumento no número de computadores no Censo Escolar. Em alguns poucos casos, foram constatados impactos negativos e significativos do Prouca. Em nenhum dos casos, foi constatado impacto significativo e positivo.

3.7 Conclusão

Neste trabalho, realizamos a primeira avaliação de impacto rigorosa do Programa Um Computador por Aluno. Como vimos, o Prouca foi implementado em 2010, quando 367 escolas receberam cerca de 150 mil *laptops*. Somente na aquisição dos *laptops*, o Governo Federal investiu quase R\$ 100 milhões. Além disso, foi investida considerável soma de recursos, financeiros e humanos, dos diversos níveis de governo envolvidos, de Universidades parcerias e das próprias escolas.

Os resultados mostraram que o programa teve impacto positivo e significativo no número de computadores por aluno nas escolas tratadas. Chama a atenção, contudo, o fato de a razão computador-aluno das escolas tratadas ter decrescido rapidamente nos anos seguintes à implementação do Programa, a ponto de quatro (cinco) anos depois, não haver qualquer diferença entre a razão computador-aluno entre as escolas tratadas e as escolas pareadas do grupo controle para os anos finais (anos iniciais) do Ensino Fundamental. O estudo mostra, assim, que houve uma rápida inutilização de computadores do programa.

O Prouca apresentou, ainda, impacto positivo e significativo no uso de computadores e da internet pelos professores como instrumento de apoio pedagógico. Esse impacto, contudo, mostrou-se de curto prazo: já em 2015, não se verifica qualquer diferença entre a utilização dessas ferramentas por professores das escolas que receberam o programa em comparação com professores de escolas que não o receberam.

As estimativas de resultado do Prouca nas taxas de abandono, no desempenho dos alunos em português e matemática e no Ideb das escolas mostram um cenário desolador para o programa. Em síntese, o programa não impactou de forma significativa nenhuma dessas variáveis em nenhum dos cinco anos seguintes a sua implementação.

Os resultados estimados no âmbito deste trabalho reforçam as conclusões da literatura acerca da inserção de TICs no ambiente escolar. Como vimos na seção 3.2, diversas intervenções de larga escala foram realizadas nos últimos anos focadas na compra e distribuição de *laptops* para escolas públicas de países em desenvolvimento. Avaliações de impacto desses programas no desempenho acadêmico dos estudantes deixam claro o surgimento de um padrão de resultados nulos, do qual, agora, o Prouca passa a fazer parte.

É importante ressaltar que os resultados aqui estimados não, necessariamente, refletem uma crítica geral ao uso da tecnologia nas escolas. Não temos dúvidas que a introdução de novas tecnologias no ambiente escolar pode contribuir para a qualidade da educação no país e, conseqüentemente, para o desempenho acadêmico dos estudantes. Contudo, evidências apresentadas neste trabalho são elucidativas no sentido de que, da forma como essa introdução foi realizada no âmbito do Prouca, a tecnologia não é capaz de contribuir, de forma significativa, para a melhora da qualidade da educação brasileira.

3.8 Referências Bibliográficas

- Alderete, M. V., & Formichella, M. M. (2017). The effect of ICTs on academic achievement: The conectar igualdad programme in Argentina. *CEPAL Review*, 2016(119), 83–100. <https://doi.org/10.18356/f23c6662-en>
- Andriola, W. B., Gomes, C. A. S., Andriola, W. B., & Gomes, C. A. S. (2017). Programa Um Computador Por Aluno (PROUCA): uma análise bibliométrica. *Educar Em Revista*, (63), 267–288. <https://doi.org/10.1590/0104-4060.48230>
- Angrist, J., & Lavy, V. (2002). New Evidence on Classroom Computers and Pupil Learning*. *The Economic Journal*, 112(482), 735–765. <https://doi.org/10.1111/1468-0297.00068>
- Bando, R., Gallego, F., Gertler, P., & Fonseca, D. R. (2017). Books or Laptops? The Effect of Shifting from Printed to Digital Delivery of Educational Content on Learning. *Economics of Education Review*. <https://doi.org/10.1016/j.econedurev.2017.07.005>
- Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remediating Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics*, 122(3), 1235–1264. <https://doi.org/10.1162/qjec.122.3.1235>
- Bebell, D., & Kay, R. (2010). One to one computing: a summary of the quantitative results from the Berkshire Wireless Learning Initiative. *Journal of Technology, Learning And Assessment*.
- Becker, S., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal*, 2(4), 358–377. <https://doi.org/The Stata Journal>
- BRASIL. (2013). *Programa Um Computador por Aluno: critérios de seleção das escolas*. Brasília.
- Bulman, G., & Fairlie, R. (2016). *Technology and Education: Computers, Software, and the Internet*. Cambridge, MA. <https://doi.org/10.3386/w22237>
- Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22(1), 31–72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>
- Card, D., & Krueger, A. (1994). Minimum Wages and Employment: A Case Study of the Fast- Food Industry in New Jersey and Pennsylvania. *The American Economic Review*, 84, 772–793.

Carrillo, P., Onofa, M., Ponce, J., Carrillo, P., Onofa, M., Ponce, J., ... Ponce, J. (2016). Information Technology and Student Achievement: Evidence from a Randomized Experiment in Ecuador. Retrieved from <https://publications.iadb.org/handle/11319/3094>

Costa, K. S. D. da. (2013). *Análise da Implantação do Programa Um Computador por Aluno no Estado do Pará*. Universidade Federal do Pará.

Cristia, J. P., Ibararán, P., Cueto, S., Santiago, A., & Severín, E. (2012). Technology and Child Development: Evidence from the One Laptop Per Child Program. Retrieved from <http://repec.iza.org/dp6401.pdf>

De Melo, G., Machado, A., & Miranda, A. (2014). The Impact of a One Laptop per Child Program on Learning: Evidence from Uruguay. *IZA Discussion Papers*. Retrieved from <https://ideas.repec.org/p/iza/izadps/dp8489.html>

Dehejia, R., & Wahba, S. (1998). Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs. Retrieved from <http://econpapers.repec.org/paper/nbrnberwo/6586.htm>

European Commission. (2013). *Survey of Schools: ICT in Education - Benchmarking Access, Use and Attitudes to Technology in Europe's Schools*.

Gomes, C. A. S. (2015). Avaliação do Programa Um computador por Aluno (PROUCA) sob a ótica do modelo CIPP. Retrieved from <http://www.repositorio.ufc.br/handle/riufc/13197>

Goolsbee, A., & Guryan, J. (2006). The Impact of Internet Subsidies in Public Schools. *Review of Economics and Statistics*, 88(2), 336–347. <https://doi.org/10.1162/rest.88.2.336>

Grimes, D., & Warschauer, M. (2008). Learning with Laptops: A Multi-Method Case Study. *Journal of Educational Computing Research*, 38(3), 305–332. <https://doi.org/10.2190/EC.38.3.d>

Hansen, N., Koudenburg, N., Hiersemann, R., Tellegen, P. J., Kocsev, M., & Postmes, T. (2012). Laptop usage affects abstract reasoning of children in the developing world. *Computers & Education*, 59(3), 989–1000. <https://doi.org/10.1016/j.compedu.2012.04.013>

Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica*, 71(4), 1161–1189. <https://doi.org/10.1111/1468-0262.00442>

- Kozma, R., McGhee, R., Quellmalz, E., & Zalles, D. (2004). Closing the digital divide: evaluation of the World Links program. *International Journal of Educational Development*, 24(4), 361–381. <https://doi.org/10.1016/j.ijedudev.2003.11.014>
- Lavinas, L., & Veiga, A. (2013). Desafios do modelo brasileiro de inclusão digital pela escola. *Cadernos de Pesquisa*, 43(149), 542–569. <https://doi.org/10.1590/S0100-15742013000200009>
- Leuven, E., Lindahl, M., Oosterbeek, H., & Webbink, D. (2007). The Effect of Extra Funding for Disadvantaged Pupils on Achievement. *Review of Economics and Statistics*, 89(4), 721–736. <https://doi.org/10.1162/rest.89.4.721>
- Linden, L. L. (2008, June 1). Complement or substitute? The effect of technology on student achievement in India. Retrieved from <http://documents.worldbank.org/curated/en/804371468034237060/Complement-or-substitute-The-effect-of-technology-on-student-achievement-in-India>
- Linden, L. L., & Barrera-Osorio, F. (2009). The use and misuse of computers in education: evidence from a randomized experiment in Colombia. Retrieved from <http://documents.worldbank.org/curated/en/346301468022433230/The-use-and-misuse-of-computers-in-education-evidence-from-a-randomized-experiment-in-Colombia>
- Machin, S., McNally, S., & Silva, O. (2006). New Technology in Schools: Is There a Payoff? *CEE Discussion Papers*. Retrieved from <https://ideas.repec.org/p/cep/ceedps/0055.html>
- Mo, D., Swinnen, J., Zhang, L., Yi, H., Qu, Q., Boswell, M., & Rozelle, S. (2013). Can One-to-One Computing Narrow the Digital Divide and the Educational Gap in China? The Case of Beijing Migrant Schools. *World Development*, 46, 14–29. <https://doi.org/10.1016/j.worlddev.2012.12.019>
- Muralidharan, K., Singh, A., & Ganimian, A. (2016). *Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India*. Cambridge, MA. <https://doi.org/10.3386/w22923>
- Nugroho, D., & Lonsdale, M. (2010). Evaluation of OLPC programs globally: a literature review. *Australian Council for Educational Research*. Retrieved from http://wiki.laptop.org/images/a/a5/OLPC_Lit_Review_v4_Aug2010.pdf
- Penuel, W. R. (2006). Implementation and Effects Of One-to-One Computing Initiatives. *Journal of Research on Technology in Education*, 38(3), 329–348. <https://doi.org/10.1080/15391523.2006.10782463>

- Pretto, N. de L., Coelho, L. A., & Almeida, L. M. T. de S. (2012). Gestão do PROUCA: a experiência do projeto piloto na Bahia. *Anais Dos Workshops Do Congresso Brasileiro de Informática Na Educação*, 1(1). <https://doi.org/10.5753/CBIE.WCBIE.2012.%P>
- Rigotto, M. E., & Souza, N. de J. de. (2006). Evolução da educação no Brasil, 1970-2003. *Análise – Revista de Administração Da PUCRS*, 16(2). Retrieved from <http://revistaseletronicas.pucrs.br/ojs/index.php/face/article/view/278>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41. <https://doi.org/10.2307/2335942>
- Santos, S. P. dos. (2014). *O programa um computador por aluno na visão dos jovens das escolas públicas de Goiânia*. Universidade Federal de Goiás. Retrieved from <https://repositorio.bc.ufg.br/tede/handle/tede/3586>
- Schneider, F. C. (2012). *Cidade Um Computador por Aluno - UCA Total Uma totalidade inclusiva em discussão*. Universidade Federal do Rio Grande do Sul.
- SEED/MEC. (2007). *UCA - Princípios Orientadores para o uso pedagógico do laptop escolar*.
- Shapley, K. D. C.-W. F. (2009). Evaluation of the Texas Technology Immersion Pilot: Final Outcomes for a Four-Year Study (2004-05 to 2007-08). *Texas Center for Educational Research*. Retrieved from <https://eric.ed.gov/?id=ED536296>
- Sharma, U., Sharma, & Uttam. (2014). Can Computers Increase Human Capital in Developing Countries? An Evaluation of Nepal's One Laptop per Child Program. Retrieved from <http://econpapers.repec.org/paper/agsaaea14/169846.htm>
- Spiezia, V. (2011). Does Computer Use Increase Educational Achievements? Student-level Evidence from PISA. *OECD Journal: Economic Studies*, 2010(1), 1–22. https://doi.org/10.1787/eco_studies-2010-5km33scwlvkf
- Sprietsma, M. (2012). Computers as pedagogical tools in Brazil: a pseudo-panel analysis. *Education Economics*, 20(1), 19–32. <https://doi.org/10.1080/09645290903546496>
- Suhr, K. A., Hernandez, D. A., Grimes, D., & Warschauer, M. (2010). Laptops and fourth-grade literacy: assisting the jump over the fourth-grade slump. *Journal of Technology, Learning, and Assessment*, 9.
- Trimmel, M., & Bachmann, J. (2004). Cognitive, social, motivational and health aspects of students in laptop classrooms. *Journal of Computer Assisted Learning*, 20(2), 151–158. <https://doi.org/10.1111/j.1365-2729.2004.00076.x>

U.S. Department of Education. (2013). *Digest of Education Statistics 2012 (NCES 2014-015)*.

UFC. (2010). *Termo de Referência - Implantação e Desenvolvimento dos Projetos-Piloto em Escolas Públicas para o Uso Pedagógico do Laptop Educacional Conectado*.

UNESCO. (2012). *ICT in Education in Latin America and the Caribbean. A Regional Analysis of ECT Integration and E-Readiness*. Retrieved from <http://www.uis.unesco.org/Communication/Documents/ict-regional-survey-lac-2012-en.pdf>

Verma, S. (2011). OLPC: Are we there yet? *Linux Journal*, 211, 98–118.

3.9 Apêndice

Tabela A.1: Variáveis de Controle do PSM

Fonte	Base	Variáveis		Descrição
		Anos Iniciais	Anos Finais	
Características Acadêmicas da Escola				
INEP	Prova Brasil (2009)	Nota_5F_Port	Nota_9F_Port	Médias das escolas em Português na Prova Brasil em 2009 (Anos Iniciais/Anos Finais)
		Nota_5F_Mat	Nota_9F_Mat	Médias das escolas em Matemática na Prova Brasil em 2009 (Anos Iniciais/Anos Finais)
	Ideb_Anos_Iniciais	Ideb_Anos_Finais	Ideb das Escolas em 2009	
	Abandono_Anos_Iniciais	Abandono_Anos_Finais	Taxa de Abandono em 2009 (anos iniciais e finais)	
	Site-INEP	Aprovação_Anos_Iniciais	Aprovação_Anos_Finais	Taxa de Aprovação em 2009 (anos iniciais e finais)
		TDI_Anos_Iniciais	TDI_Anos_Finais	Taxa de distorcao idade-serie em 2009 (anos iniciais e finais)
Características da Escola (Infraestrutura, Tamanho e Etapas de Ensino)				
INEP	Censo Escolar (2009)	id_dependencia_adm	id_dependencia_adm	variável dummy, igual a um para escolas estaduais
		id_internet	id_internet	variável dummy, igual a um se as escolas possuírem internet
		id_banda_larga	id_banda_larga	variável dummy, igual a um se as escolas possuírem internet banda larga
		id_biblioteca	id_biblioteca	variável dummy, igual a um se a escola possuir uma biblioteca/sala de leitura
		id_quadra_esportes	id_quadra_esportes	variável dummy, igual a um se a escola possuir uma quadra de esportes
		num_salas_existentes	num_salas_existentes	número de salas de aula disponíveis na escola
		num_computadores	num_computadores	quantidade de Computadores disponíveis na escola

	computador_por_aluno	computador_por_aluno	quantidade de computadores disponíveis por aluno na escola
	id_laboratorio_ciencias	id_laboratorio_ciencias	variável dummy, igual a um se a escola possuir um Laboratório de Ciências
	id_laboratorio_informatica	id_laboratorio_informatica	variável dummy, igual a um se a escola possuir um Laboratório de Informática
	id_sala_diretoria	id_sala_diretoria	variável dummy, igual a um se a escola possuir sala para a diretoria
	id_sala_professor	id_sala_professor	variável dummy, igual a um se a escola possuir salas de professores
	id_equip_impresora	id_equip_impresora	variável dummy, igual a um se a escola possuir uma impressora
	num_funcionarios	num_funcionarios	número de funcionários que trabalham na escola (inclusive professores e auxiliares)
	numero_de_alunos	numero_de_alunos	número de alunos matriculados no Ensino Regular da Escola em 2009
	tem_ef_finais	tem_ef_iniciais	variável dummy, igual a um se a escola possui anos iniciais/anos finais do Ensino Fundamental
	tem_em	tem_em	variável dummy, igual a um se a escola possui Ensino Médio
	tem_ed_infantil	tem_ed_infantil	variável dummy, igual a um se a escola possui Educação Infantil
<hr/>			
Participação em outros programas do Governo Federal			
	pnbl	pnbl	Escola conectada pelo programa banda larga na escola entre 2008 e 2010
	gesac	gesac	Escola conectada pelo programa gesac entre 2008 e 2010
Ministério da Educação	Ministério da Educação	proinfo	Escola com laboratório entregue pelo proinfo entre 2006 e 2010
		pme	Escola participante do Programa Mais Educação (2008 e/ou 2009)

Características do Prefeito

		partido_pt_2009	partido_pt_2009	Prefeito filiado ao PT em 2009
		partido_oposicao_2009	partido_oposicao_2009	Prefeito filiado ao PSDB, DEM ou PPS em 2009
IBGE	IBGE	alinhamento_governador	alinhamento_governador	Prefeito alinhado com o governador o estado em 2009
		prefeito_ensino_superior	prefeito_ensino_superior	Prefeito possui ensino superior

Características dos Municípios

IBGE	IBGE	pib_pcap populacao	pib_pcap populacao	pib per capita em 2009 populacao em 2010
PNUD	PNUD	idhm	idhm	Índice de Desenvolvimento Humano Municipal (2010)

Tabela A.2: Estatísticas Descritivas e testes de médias – Anos Finais

	Escolas Tratadas = 95		Escolas Controle = 21.744		Teste de Igualdade de Médias	
	Média	Desvio padrão	Média	Desvio padrão	Diferença	p-valor
Prova Brasil						
Nota_9F_Port	241.78	19.00	238.35	18.61	3.43	0.07
Nota_9F_Mat	244.67	20.69	241.84	19.82	2.82	0.17
Ideb_Anos_Finais	3.86	0.86	3.80	0.83	0.06	0.46
Abandono_Anos_Finais	5.17	5.93	4.92	5.68	0.25	0.67
Aprovação_Anos_Finais	80.38	12.38	81.11	11.90	-0.73	0.55
TDI_Anos_Finais	30.41	13.88	29.60	14.51	0.81	0.59
Censo Escolar 2009						
id_dependencia_adm	0.60	0.49	0.59	0.49	0.01	0.86
id_internet	0.85	0.36	0.84	0.37	0.01	0.70
id_banda_larga	0.71	0.46	0.67	0.47	0.04	0.46
id_biblioteca	0.71	0.46	0.66	0.47	0.04	0.36
id_quadra_esportes	0.68	0.47	0.73	0.44	-0.05	0.27
num_salas_existentes	10.19	4.29	12.83	5.70	-2.64	0.00
num_computadores2009	22.05	48.79	16.26	12.18	5.79	0.00
computador_por_aluno2009	0.05	0.07	0.02	0.02	0.02	0.00
id_laboratorio_ciencias	0.14	0.35	0.28	0.45	-0.14	0.00
id_laboratorio_informatica	0.84	0.37	0.81	0.39	0.03	0.41
id_sala_diretoria	0.91	0.29	0.96	0.20	-0.05	0.01
id_sala_professor	0.92	0.28	0.94	0.25	-0.02	0.44
id_equip_impresora	0.99	0.10	0.97	0.18	0.02	0.20

num_funcionarios	47.63	22.61	62.15	29.32	-14.52	0.00
numero_de_alunos2009	506.52	295.28	788.62	430.33	-282.10	0.00
participantes_5F_2009	46.56	30.35	72.04	46.90	-25.48	0.00
tem_ef_iniciais	0.74	0.44	0.61	0.49	0.13	0.01
tem_em	0.28	0.45	0.46	0.50	-0.17	0.00
tem_ed_infantil	0.20	0.40	0.15	0.36	0.05	0.16
<hr/>						
Dados_MEC						
<hr/>						
proinfo	0.73	0.45	0.68	0.47	0.05	0.33
gesac	0.15	0.36	0.08	0.27	0.07	0.02
pnbl	0.78	0.42	0.71	0.45	0.07	0.14
pme	0.07	0.26	0.09	0.29	-0.02	0.49
<hr/>						
IBGE_Prefeitos						
<hr/>						
partido_pt_2009	0.13	0.33	0.16	0.36	-0.03	0.42
partido_oposicao_2009	0.27	0.45	0.28	0.45	0.00	0.92
prefeito_ensino_superior	0.64	0.48	0.67	0.47	-0.03	0.51
sexo_prefeito	0.94	0.24	0.90	0.29	0.03	0.28
alinhamento_governador	0.26	0.44	0.20	0.40	0.06	0.15
<hr/>						
IBGE_Municipios						
<hr/>						
pib_pcap2009	16,004	11,834	16,118	13,959	-114.24	0.94
populacao2009	737,147	2,244,774	884,696	2,418,965	-147,548.90	0.55
idhm	0.71	0.08	0.71	0.08	0.00	0.53
<hr/>						

Fonte: Elaboração própria a partir dos dados pesquisados.

Tabela A.3: Balanceamento do Escore de Propensão (método do vizinho mais próximo) – Anos Iniciais

Variável	Matching	Média		Viés		p>t
		Tratamento	Controle	% viés	% de redução	
Prova Brasil						
Nota_5F_Port	Antes	179.65	182.32	-13.30		0.12
	Depois	179.65	181.24	-8.0	40.3	0.48
Nota_5F_Mat	Antes	199.33	202.66	-14.3		0.10
	Depois	199.33	200.78	-6.2	56.6	0.58
Ideb_Anos_Iniciais	Antes	4.44	4.59	-14.7		0.09
	Depois	4.44	4.48	-4.1	72.5	0.72
Abandono_Anos_Iniciais	Antes	1.85	1.91	-1.9		0.82
	Depois	1.85	2.10	-7.8	-319.3	0.51
Aprovação_Anos_Iniciais	Antes	88.03	88.68	-7.3		0.38
	Depois	88.03	87.82	2.4	67.1	0.84
TDI_Anos_Iniciais	Antes	19.47	18.27	10.1		0.24
	Depois	19.47	19.82	-3.0	70.8	0.80
Censo Escolar 2009						
id_dependencia_adm	Antes	0.30	0.24	11.9		0.13
	Depois	0.30	0.32	-4.5	62.0	0.71
id_internet	Antes	0.76	0.74	5.2		0.53
	Depois	0.76	0.70	12.3	-137.6	0.30
id_banda_larga	Antes	0.55	0.58	-6.1		0.46
	Depois	0.55	0.53	4.1	33.5	0.73
id_biblioteca	Antes	0.66	0.57	18.8		0.03
	Depois	0.66	0.68	-5.5	70.6	0.62
id_quadra_esportes	Antes	0.51	0.55	-8.7		0.29
	Depois	0.51	0.56	-9.4	-7.8	0.42
num_salas_existentes	Antes	10.18	11.05	-16.5		0.06
	Depois	10.18	10.17	0.3	98.4	0.98
num_computadores	Antes	15.75	10.82	16.9		0.00
	Depois	15.75	13.68	7.1	58.0	0.58
computador_por_aluno	Antes	0.03	0.02	30.7		0.00
	Depois	0.03	0.03	5.1	83.4	0.72
id_laboratorio_ciencias	Antes	0.09	0.10	-4.2		0.62
	Depois	0.09	0.10	-4.6	-9.2	0.69
id_laboratorio_informatica	Antes	0.66	0.58	17.0		0.04
	Depois	0.66	0.63	5.5	67.4	0.63
id_sala_diretoria	Antes	0.89	0.93	-13.2		0.07
	Depois	0.89	0.93	-11.8	10.5	0.32
id_sala_professor	Antes	0.90	0.84	16.8		0.06

	Depois	0.90	0.90	0.0	100.0	1.00
id_equip_imprensa	Antes	0.93	0.93	1.2		0.88
	Depois	0.93	0.95	-5.3	-339.5	0.63
num_funcionarios	Antes	47.21	48.10	-3.6		0.69
	Depois	47.21	46.95	1.1	70.6	0.92
numero_de_alunos	Antes	469.87	576.66	-34.0		0.00
	Depois	469.87	476.93	-2.3	93.4	0.82
participantes_5F	Antes	51.40	69.70	-50.1		0.00
	Depois	51.40	49.67	4.7	90.6	0.57
tem_ef_finais	Antes	0.68	0.53	31.0		0.00
	Depois	0.68	0.74	-11.1	64.1	0.31
tem_em	Antes	0.05	0.09	-14.8		0.11
	Depois	0.05	0.03	10.3	30.2	0.24
tem_ed_infantil	Antes	0.42	0.39	5.4		0.51
	Depois	0.42	0.42	-1.4	74.6	0.91
<hr/>						
Dados_MEC						
proinfo	Antes	0.72	0.68	9.2		0.27
	Depois	0.72	0.71	1.5	84.1	0.90
gesac	Antes	0.09	0.04	20.5		0.00
	Depois	0.09	0.06	11.1	45.8	0.38
pnbl	Antes	0.83	0.79	11.6		0.18
	Depois	0.83	0.87	-8.5	26.4	0.42
pme	Antes	0.09	0.10	-3.1		0.71
	Depois	0.09	0.12	-9.0	-185.5	0.46
<hr/>						
IBGE_Prefeitos						
partido_pt_2009	Antes	0.15	0.16	-0.9		0.92
	Depois	0.15	0.12	9.2	-961.1	0.40
partido_oposicao_2009	Antes	0.17	0.25	-19.8		0.02
	Depois	0.17	0.15	5.0	74.8	0.64
prefeito_ensino_superior	Antes	0.57	0.67	-19.9		0.01
	Depois	0.57	0.56	2.8	86.1	0.82
alinhamento_governador	Antes	0.23	0.21	4.9		0.54
	Depois	0.23	0.21	6.4	-30.3	0.58
<hr/>						
IBGE_Municipios						
pib_pcap2009	Antes	14114	15758	-13.3		0.17
	Depois	14114	13993	1.0	92.6	0.91
populacao2009	Antes	350000	650000	-20.3		0.04
	Depois	350000	370000	-0.8	95.9	0.91
idhm2010	Antes	0.70	0.71	-13.9		0.09
	Depois	0.70	0.70	0.70	0.70	0.70

Fonte: Elaboração própria a partir dos dados pesquisados.

Tabela A.4: Balanceamento do Escore de Propensão (método do vizinho mais próximo) – Anos Finais

Variável	Matching	Média		Viés		p>t
		Tratamento	Controle	% viés	% de redução	
Prova Brasil						
Nota_9F_Port	Antes	241.8	238.4	18.2		0.07
	Depois	241.8	242.7	-4.7	74.1	0.75
Nota_9F_Mat	Antes	244.7	241.8	13.9		0.17
	Depois	244.7	245.8	-5.6	59.9	0.71
Ideb_Anos_Finais	Antes	3.86	3.80	7.4		0.46
	Depois	3.86	3.89	-4	46.3	0.79
Abandono_Anos_Finais	Antes	5.17	4.92	4.3		0.67
	Depois	5.17	4.99	3.1	28.1	0.84
Aprovação_Anos_Finais	Antes	80.38	81.11	-6.0		0.55
	Depois	80.38	80.43	-0.4	93.8	0.98
TDI_Anos_Finais	Antes	30.42	29.60	5.7		0.59
	Depois	30.42	30.12	2.1	63.3	0.89
Censo Escolar 2009						
id_dependencia_adm	Antes	0.60	0.59	1.8		0.86
	Depois	0.60	0.59	2.1	-20.2	0.88
id_internet	Antes	0.85	0.84	4.1		0.70
	Depois	0.85	0.93	-20.3	-396.6	0.11
id_banda_larga	Antes	0.71	0.67	7.6		0.46
	Depois	0.71	0.73	-4.5	40.7	0.75
id_biblioteca	Antes	0.71	0.66	9.6		0.36
	Depois	0.71	0.73	-4.5	53.1	0.75
id_quadra_esportes	Antes	0.68	0.73	-11.0		0.27
	Depois	0.68	0.66	4.6	57.9	0.76
num_salas_existentes	Antes	10.19	12.83	-52.4		0.00
	Depois	10.19	10.22	-0.6	98.8	0.96
num_computadores	Antes	22.05	16.26	16.3		0.00
	Depois	22.05	20.17	5.3	67.5	0.80
computador_por_aluno	Antes	0.05	0.02	43.9		0.00
	Depois	0.05	0.04	5.4	87.7	0.79
id_laboratorio_ciencias	Antes	0.14	0.28	-35.7		0.00
	Depois	0.14	0.19	-13.1	63.2	0.33
id_laboratorio_informatica	Antes	0.84	0.81	8.7		0.41
	Depois	0.84	0.82	5.5	36.4	0.70
id_sala_diretoria	Antes	0.91	0.96	-20.3		0.01
	Depois	0.91	0.82	33.3	-63.9	0.09
id_sala_professor	Antes	0.92	0.94	-7.5		0.44

	Depois	0.92	0.86	20	-167.3	0.25
id_equip_imprensa	Antes	0.99	0.97	16.3		0.20
	Depois	0.99	0.99	0	100	1.00
num_funcionarios	Antes	47.63	62.15	-55.4		0.00
	Depois	47.63	46.98	2.5	95.5	0.83
numero_de_alunos	Antes	506.52	788.6	-76.4		0.00
	Depois	506.52	494.1	3.4	95.6	0.75
participantes_9F	Antes	46.56	72.04	-64.5		0.00
	Depois	46.56	47.28	-1.8	97.1	0.87
tem_ef_finais	Antes	0.74	0.61	27.3		0.01
	Depois	0.74	0.73	2.3	91.7	0.87
tem_em	Antes	0.28	0.46	-36.5		0.00
	Depois	0.28	0.28	0	100	1.00
tem_ed_infantil	Antes	0.20	0.15	13.6		0.16
	Depois	0.20	0.23	-8.3	38.9	0.60
<hr/>						
Dados_MEC						
<hr/>						
proinfo	Antes	0.73	0.68	10.1		0.33
	Depois	0.73	0.73	0	100	1.00
gesac	Antes	0.15	0.08	21.2		0.02
	Depois	0.15	0.16	-3.3	84.3	0.84
pnbl	Antes	0.78	0.71	15.9		0.14
	Depois	0.78	0.73	12.1	24.3	0.40
pme	Antes	0.07	0.09	-7.5		0.49
	Depois	0.07	0.03	15.1	-102.2	0.20
<hr/>						
IBGE_Prefeitos						
<hr/>						
partido_pt_2009	Antes	0.13	0.16	-8.6		0.42
	Depois	0.13	0.14	-3	64.9	0.83
partido_oposicao_2009	Antes	0.27	0.28	-1.0		0.92
	Depois	0.27	0.27	0	100	1.00
prefeito_ensino_superior	Antes	0.64	0.67	-6.7		0.51
	Depois	0.64	0.62	4.4	33.8	0.77
alinhamento_governador	Antes	0.26	0.20	14.0		0.15
	Depois	0.26	0.26	0	100	1.00
<hr/>						
IBGE_Municipios						
<hr/>						
pib_pcap2009	Antes	16004	16118	-0.9		0.94
	Depois	16004	15084	7.1	-705.2	0.59
populacao2009	Antes	740000	880000	-6.3		0.55
	Depois	740000	840000	-4.4	30.3	0.76
idhm	Antes	0.71	0.71	-6.2		0.53
	Depois	0.71	0.70	3	51.7	0.84

Fonte: Elaboração própria a partir dos dados pesquisados.