

Impacto Sobre o Desempenho Escolar do Pagamento de Bônus aos Docentes do Ensino Fundamental do Estado de São Paulo^{*,†}

Cláudia Hiromi Oshiro[‡], Luiz Guilherme Scorzafave[§], Tulio Anselmi Dorigan[¶]

Sumário: 1. Introdução; 2. Revisão literária; 3. O programa de bonificação de São Paulo; 4. Metodologia; 5. Dados; 6. Resultados; 7. Análise de robustez; 8. Considerações finais; A. Apêndice.

Palavras-chave: Educação, Bônus, Accountability, Salário, Pagamento a Professores.

Códigos JEL: I25, I28.

Políticas de bonificação têm se popularizado como mecanismo de indução de melhoria da qualidade da educação. Dada a escassa evidência dessas políticas para o Brasil, este artigo avalia o efeito do programa de bonificação de professores do estado de São Paulo sobre o desempenho escolar dos alunos utilizando os métodos de *Propensity Score Matching* e Diferenças em Diferenças. Foram encontrados efeitos positivos e significativos para Matemática (0,42 desvios-padrão) e Português (0,14 desvios-padrão) para o quinto ano do ensino fundamental entre 2007 e 2009, mas em 2011 há uma diminuição expressiva desse efeito positivo. Para o nono ano, não foram encontrados efeitos e, para algumas especificações e grupos de controle utilizados, houve impacto negativo sobre a proficiência em 2009 e 2011. Os resultados foram robustos a diferentes grupos de controle.

Bonus payment schemes to teachers have been growing overseas as a tool for education quality improvement. Since there is very little evidence concerning this kind of policy in Brazil, this paper aims to evaluate the effect of bonus program for teachers in the state of Sao Paulo, Brazil, over proficiency of students. Propensity Score Matching and Difference in Differences methods are used to measure the impact of the program. We found positive and significant effects on Math tests (0,42 standard-deviations) and Lecture tests (0,14 standard-deviations) for the fourth grade between

*Os autores agradecem aos comentários dos pareceristas anônimos, bem como os de Lígia Vasconcellos, Priscila Albuquerque, Miguel Foguel, Cláudio Ferraz, Rafael Camelo, Cristine Xavier Pinto e Elaine Toldo Pazello.

†Endereço para correspondência: Luiz Guilherme Scorzafave. FEA. Av. Bandeirantes, 3900, Ribeirão Preto-SP. CEP 14040-905.

‡Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas (FGV/EAESP). Doutoranda em Administração Pública e Governo. E-mail: claudiahiromi@gmail.com

§Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto / Universidade de São Paulo (FEA-RP/USP). E-mail: scorza@usp.br

¶Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto / Universidade de São Paulo (FEA-RP/USP). E-mail: tulioandorigan@gmail.com



2007 and 2009, but there is a sharp decay in these effects in 2011. We didn't find any effects for eighth grade of elementary school and, in some specifications and control groups, the results were negative in 2009 and 2011. These results are robust to different control groups.

1. INTRODUÇÃO

A melhoria do sistema educacional vem se mostrando um grande desafio para a sociedade brasileira nos últimos anos. A quase universalização do ensino fundamental nas últimas décadas representou um grande avanço para a educação. Porém, a formação de capital humano requer educação de qualidade e, nesse sentido, o sistema brasileiro ainda deixa a desejar. O nível educacional brasileiro ainda é muito baixo quando comparado a outros países. Segundo o Programa Internacional de Avaliação de Alunos (Pisa) de 2009, o Brasil ocupou a 54ª posição entre 65 países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE).

As políticas públicas educacionais em todo o mundo têm voltado suas atenções para a melhora da qualidade em todos os níveis de formação escolar. Diante desse desafio, as reformas na área têm como um dos temas centrais programas de responsabilização dos agentes escolares pela melhoria do desempenho dos alunos.

No Brasil, a expansão das avaliações externas do ensino a partir da década de 1990 tornou possível a elaboração de mecanismos de responsabilização. Um importante passo foi a reformulação do Sistema Nacional de Avaliação da Educação Básica (Saeb), em 1995. Com o passar dos anos, houve uma evolução considerável dos sistemas de avaliação nas esferas estaduais e municipais.

A divulgação pública dos resultados dos exames mais recentes funciona como uma forma de responsabilizar as escolas e seus funcionários pelo desempenho dos alunos e, desta forma, criar incentivos para melhoria. O conhecimento público das notas pode gerar pressão por parte dos pais e da comunidade como um todo sobre as instituições com piores resultados. Também pode haver premiações ou punições condicionadas ao desempenho em exames.

Diante desse contexto, o objetivo deste trabalho é investigar se o programa de bonificação de professores e funcionários da Secretaria de Educação do Estado de São Paulo, implantado em 2008, ao responsabilizar os agentes escolares pelo desempenho dos alunos, tem algum efeito sobre as notas nos testes padronizados. A ideia desse tipo de política é que ao pagar uma bonificação aos docentes, haveria um incentivo para que os mesmos mudassem suas práticas escolares, realizando um maior esforço e mostrando mais comprometimento com a melhoria do aprendizado dos alunos.

Para este estudo, foi utilizada a metodologia de *Propensity Score Matching* e de Diferenças em Diferenças. Esta abordagem consiste, basicamente, em encontrar um grupo de controle mais parecido com o grupo tratamento com base em características observáveis e, então, comparar os resultados ao longo do tempo, para eliminar diferenças não observáveis fixas no tempo entre as escolas afetadas pelo tratamento, ou seja, as estaduais paulistas e as não afetadas. Para isso, são utilizados dados da Prova Brasil de 2005 a 2011. Diversas análises de robustez são levadas a cabo para assegurar a validade dos resultados.

Além da introdução, em seguida é apresentada uma breve revisão da literatura pertinente ao tema. Na [seção 3](#) o programa de bonificação paulista é detalhado, e na [seção 4](#) é apresentada a metodologia empregada. A [seção 5](#) apresenta a descrição da base de dados e algumas estatísticas descritivas. Na [seção 6](#) encontram-se os resultados da análise econométrica, que são validados a partir dos testes de robustez contidos na [seção 7](#). Por fim, a [última seção](#) apresenta as considerações finais do trabalho.

2. REVISÃO LITERÁRIA

Uma das características das políticas educacionais mais recentes tem sido a busca de alternativas à mera distribuição de recursos e insumos escolares da forma como é tradicionalmente realizada, uma

vez que essas práticas vêm se mostrando ineficientes, já que o aumento dos gastos com educação, por si só, não leva necessariamente a um crescimento da qualidade do ensino. [Hanushek \(2003\)](#) mostrou que recursos escolares estão pouco relacionados com o desempenho dos alunos e com a variação da qualidade dos professores.

Dentro do contexto de valorização da profissão de docente, vários países têm adotado políticas de pagamento por mérito. Entretanto, os resultados desse tipo de política não são consensuais. [Holmstrom e Milgrom \(1991\)](#), ao analisarem o problema do agente-principal sob modelo multitarefas, citam as discussões sobre os incentivos na área de educação. Os defensores dessa prática se baseiam na teoria microeconômica do problema unidimensional padrão de agente-principal, no qual o incentivo levaria os professores a um maior esforço com impacto positivo sobre a proficiência do aluno. Já o contra-argumento é que o incentivo levaria o professor a colocar maior esforço em tarefas que podem ser medidas por meio de exames padronizados, deixando de lado funções importantes como, por exemplo, o desenvolvimento da criatividade e de diversas habilidades não cognitivas que talvez não sejam medidas nas provas.

[Vigdor \(2008\)](#) investigou o programa de bonificação de professores da Carolina do Norte — um dos mais antigos dos Estados Unidos — utilizando dados de 2002/03 e 2005/06 e observou melhora do desempenho nos exames de Matemática. Porém, o autor faz ressalvas quanto à estrutura de incentivos adotada — baseada nos resultados gerais da escola —, que pode levar ao problema do *free-rider*. O mesmo programa foi avaliado por [Lauen \(2011\)](#), mas para um período diferente. Desta vez, foram encontrados efeitos positivos e significativos apenas para o exame de Leitura do quinto ano (ganho de 0,1 desvio padrão).

[Figlio e Kenny \(2007\)](#) combinaram dados do *National Education Longitudinal Survey*, de 2000, com uma pesquisa própria e fizeram uma análise cross-section para verificar a relação entre incentivos aos professores e desempenho dos alunos nos Estados Unidos. Segundo eles, escolas que oferecem pagamento por mérito apresentaram notas maiores, sobretudo aquelas que têm um mínimo de participação dos pais. Entretanto, não descartam a possibilidade de que escolas mais eficientes — e que, portanto, tem pontuação maior — sejam mais propensas a adotar essas políticas.

[Lavy \(2009\)](#) avaliou um programa experimental em Israel que recompensava individualmente os professores — com bônus em dinheiro — pela melhora do desempenho dos estudantes. O desenho do programa consistia em um torneio, no qual o pagamento dependia da posição relativa do professor entre aqueles de mesma área de atuação. Os resultados indicam que houve melhora do desempenho dos estudantes em exames durante o ensino médio utilizados como requisito para ingresso no ensino superior.

Também em caráter temporário, [Glewwe, Ilias e Kremer \(2003\)](#) avaliaram um programa de premiação de professores no Quênia condicionado ao desempenho dos estudantes e encontraram maiores ganhos de notas em escolas participantes durante o período do programa. Todavia, não verificaram mudanças em atitudes dos professores, seja em relação à maior presença em sala de aula ou nos métodos pedagógicos, a não ser na alocação de tempo para treinamento para a prova, o que pode explicar o aumento das notas. [Muralidharan e Sundararaman \(2011\)](#) analisaram os resultados de um experimento aleatório no estado rural indiano de Andhra Pradesh que bonificava professores de escolas primárias que obtivessem melhora da pontuação média, e os resultados encontrados após dois anos de programa foram de aumento de 0,27 e 0,17 desvio padrão em testes de Matemática e de Linguagem, respectivamente.

Entre os estudos que não mostraram efeito positivos, [Glazerman, Carey e McKie \(2009\)](#) avaliaram os efeitos do *Teacher Advancement Program* (TAP) de Chicago, implementado em 2007, que promove uma remuneração extra e novas responsabilidades ao docente por meio da designação de Professor Mentor. A bonificação fica atrelada ao desempenho do aluno em sala de aula. Não foram encontrados sinais de melhora com relação às notas ou à satisfação dos professores. [Springer et al. \(2010\)](#) analisaram os resultados do *Project on Incentives in Teaching* (POINT), um projeto realizado no Sistema Escolar Metropolitano de Nashville, entre 2006/07 e 2008/09, do qual participavam voluntariamente professores de Matemática do *middle school* (sexta a oitava séries) que eram remunerados de acordo com ganhos de notas dos alunos em testes padronizados. Também não foram encontrados resultados muito favoráveis. Os alunos



dos professores do grupo de tratamento não tiveram desempenho melhor do que aqueles cujos professores estavam no grupo controle. Evidências positivas foram encontradas apenas para a quinta série, sobretudo no segundo ano do experimento. Fryer (2013) analisou o impacto dos incentivos financeiros a professores sobre o desempenho dos alunos na cidade de Nova Iorque no período de 2006/2007 a 2009/2010. Além de buscar os efeitos sobre a proficiência, o estudo verificou se houve alguma alteração sobre o comportamento do professor por meio de análise de dois indicadores: falta e repetência. Não foram encontradas evidências de mudança no comportamento dos professores ou alunos, tampouco indício de melhora da proficiência. Pelo contrário, os efeitos estimados foram negativos para ensino básico e *middle school*.

Com relação à literatura brasileira, poucos são os estudos empíricos na área. Apesar dos grandes avanços em termos de mobilização social e em debates sobre a educação, o que dá força para maior responsabilização dos agentes escolares, praticamente não há tentativas de avaliação dos programas de *accountability* forte brasileiros. Isso torna ainda mais relevante o presente artigo, ainda mais diante da evidência internacional inconclusiva acerca da eficácia de tais políticas.¹

3. O PROGRAMA DE BONIFICAÇÃO DE SÃO PAULO

As regras do sistema de remuneração paulista foram lançadas pelo Programa de Qualidade da Escola (PQE) em maio de 2008 (Governo do Estado de São Paulo/SEE, 2009b) e o projeto de lei referente ao programa foi aprovado apenas em dezembro do mesmo ano². A Secretaria de Educação do Estado de São Paulo estipulou metas para cada série da escola e remunera todos professores do ciclo avaliado e os demais funcionários. As metas são estabelecidas com base no Índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp) para cada etapa escolar s . O Idesp, por sua vez, é composto por dois outros índices:

$$\text{IDESP}_s = \text{ID}_s \times \text{IF}_s. \quad (1)$$

ID é o Índice de Desempenho e IF é o Índice de Fluxo. O ID é baseado nas notas do Saresp, exame de avaliação externa paulista, que é aplicado para a segunda, quarta, sexta e oitava séries (ou terceiro, quinto, sétimo e nono ano) do ensino fundamental, além da terceira série do ensino médio nas escolas administradas pelo governo estadual.³ O índice representa a quantidade relativa de alunos em cada nível de proficiência, divididas em quatro níveis: (i) Abaixo do básico, (ii) Básico, (iii) Adequado, e (iv) Avançado. O Saresp avalia a proficiência dos alunos de acordo com as exigências do Sistema Nacional de Avaliação da Educação Básica, que define o aprendizado desejado para o aluno no final de cada ciclo. Para o pagamento de bônus, considera-se apenas a proficiência em Matemática e Português.

A partir dos dados do IDESP, calcula-se a parcela de cumprimento da meta, por meio do Índice de Cumprimento (IC) em determinado ano t para cada etapa escolar s , por meio da seguinte expressão:

$$\text{IC} = \left(\frac{\text{IDESP}_{s,t} - \text{IDESP}_{s,t-1}}{\text{META}_{s,t} - \text{IDESP}_{s,t-1}} \right). \quad (2)$$

Ou seja, o IC é dado pela razão da variação do IDESP entre dois anos no ano pela variação da meta da escola na série s no ano t , em relação ao IDESP do ano.⁴ O IC é limitado no intervalo de 0% a 120%. Para

¹ Cabe destacar que estão em andamento pelo menos dois outros trabalhos avaliando o impacto do sistema de bonificação dos docentes no desempenho escolar no Brasil: Ferraz e Bruns (2012) avaliando a experiência de Pernambuco e Guerra e Foguel (2012), comparando os casos de São Paulo e Pernambuco.

² Projeto de Lei Complementar nº 1.078.

³ Como o período analisado corresponde exatamente à transição do ensino fundamental de oito para o de nove anos, ao longo do texto utilizou-se quarta série e quinto ano de modo intercambiável. Idem para oitava série e nono ano.

⁴ As metas anuais foram estabelecidas para cada escola de acordo a distância da meta de longo prazo, seguindo uma função de crescimento logística (Ferraz, 2009). Desta forma, escolas que partem de uma pontuação baixa têm que apresentar uma melhora porcentual maior com relação àquelas com maiores notas.

valores negativos, considera-se o piso (zero). Para valores superiores a 120%, considera-se o cumprimento de 120%.

Um adicional por qualidade passou a ser considerado a partir de 2009. Este índice reflete a posição da escola em relação à média das outras escolas ($IDESP^{GLOBAL}$) e à sua meta de longo prazo (2030). O Índice de Qualidade (IQ) do ano t para a série s é calculado pela seguinte fórmula:

$$IQ_{s,t} = \left(\frac{IDESP_{s,t} - IDESP_{s,t}^{GLOBAL}}{META_{s,2030} - IDESP_{s,t}^{GLOBAL}} \right). \quad (3)$$

As metas de longo prazo do IDESP para cada série são iguais para todas as escolas ($META_{s,2030}$) e foram estabelecidas levando-se em consideração níveis de aprendizado comparáveis aos de países da Organização para a Cooperação e o Desenvolvimento Econômico (OCDE), correspondendo a 7,0 para o 5º ano e 6,0 para o 9º ano do ensino fundamental.

Para efeito de pagamento do bônus, o sistema paulista considera a soma dos indicadores (IQ + IC), limitado entre 0 e 120%. O valor pago varia de acordo com a parcela da meta cumprida e é feito a professores do ciclo que cumpriram a meta e para funcionários com nível de absenteísmo abaixo de 1/3. Em escolas que cumprem 100% da meta, por exemplo, os funcionários recebem o equivalente a 2,4 salários. Em escolas que superam a meta prevista, o bônus pode chegar a 2,9 salários. O número de salários recebidos como bônus é proporcional a percentagem de atendimento da meta. Por exemplo, se a escola alcançou 50% da meta, os funcionários recebem 1,2 salários de bônus. Escolas que não cumprem as metas, mas que, nos dois anos anteriores foram classificadas entre as 10% melhores, também recebem uma remuneração. Esse seria um incentivo para escolas boas, cujas metas são difíceis de cumprir, pois as notas já estariam num patamar relativamente elevado. Funcionários das diretorias regionais recebem bônus pela média da região em que atuam.

A Tabela 1 apresenta os dados a respeito da abrangência e das características básicas do sistema de bonificação paulista. Os pagamentos são feitos no início do ano seguinte ao ano-referência para cálculo do bônus. Com exceção de 2010, que teve um valor menor de bônus médio pago, nos demais anos o bônus médio variou entre R\$ 2.600,00 e R\$ 3.100,00, o que equivalia a cerca de 2,5 salários até 2010 e a 1,9 salários mensais em 2012. Ou seja, o bônus médio equivale ao recebimento de 14º e 15º salários, um

Tabela 1. Caracterização dos pagamentos do sistema de bonificação paulista.

Ano	Data do pagamento	Valor total pago (R\$ milhões)	Nº de beneficiados	Bônus médio (R\$)*	Nº salários médio*
2008	31/03/2009	590,6	195.504	3.020,91	2,5
2009	25/03/2010	655,0	210.000	3.119,05	2,5
2010	01/04/2011	340,0	190.000	1.789,47	2,4
2011	30/03/2012	538,5	205.329	2.622,62	2,1
2012	28/03/2012	590,2	206.000	2.865,05	1,9

* Calculado excluindo-se escolas que alcançaram 0% da meta.

Fonte: Elaboração própria a partir da consulta os seguintes sites:

<http://www.educacao.sp.gov.br/noticias/professores-do-estado-recebem-hoje-pela-primeira-vez-bonus-por-desempenho>

<http://www.educacao.sp.gov.br/noticias/educacao-paga-r-655-milhoes-em-bonus-por-resultado>

<http://www.educacao.sp.gov.br/noticias/educacao-paga-r-340-milhoes-em-bonus>

<http://www.educacao.sp.gov.br/noticias/bonus-por-desempenho-da-rede-estadual-de-ensino-aumenta-59-2>

<http://www.educacao.sp.gov.br/noticias/numero-de-servidores-estaduais-que-ganharao-bonus-por-desempenho-acima-de-r-5-mil-aumenta-25>



valor em linha com alguns sistemas de bonificação americanos. Por exemplo, [Vigdor \(2008\)](#) mostra que o programa da Carolina do Norte pagava de US\$ 750,00 a US\$ 1.500,00 por docente, valor semelhante ao adotado em São Paulo. [Fryer \(2013\)](#) mostra que o bônus de Nova York é mais generoso (US\$ 3.000,00 por docente sindicalizado, a ser dividido como a escola preferir, se alcance das metas for de 100%).

Uma outra característica importante dos sistemas de bonificação é a dificuldade (ou não) para o alcance das metas. Se o sistema for desenhado de tal sorte que as metas sejam facilmente alcançadas, a lógica da bonificação de premiar um maior esforço na melhoria da prática docente fica prejudicado. Por outro lado, se a meta for considerada muito difícil de ser alcançada, também pode desestimular a mudança de comportamento por parte dos agentes envolvidos no processo.

Nesse sentido, a [Tabela 2](#) apresenta a proporção de escolas que não receberam bônus em cada ano, e, para aquelas que ganharam algum bônus, qual a porcentagem média de alcance das metas. Aqui ficam evidentes as diferenças entre escolas de 5º e 9º ano. Desde 2009, todas as escolas têm alcançado, em média, uma fração menor da meta; porém, a queda foi muito mais pronunciada no 9º ano, especialmente a partir de 2011. Por sua vez, a fração de escolas que não receberam bônus varia bastante ao longo dos anos (entre 14% e 49%), com comportamento mais parecido entre 5º e 9º ano, exceto em 2010. Ou seja, o sistema paulista não está nem entre os mais rígidos, nos quais praticamente todos os docentes são beneficiados, nem entre os menos rigorosos, onde praticamente ninguém consegue receber a bonificação.

Por fim, a [Figura A-1](#) do [Apêndice A](#) mostra a distribuição das escolas por cumprimento das metas. Em geral, as distribuições são bimodais em torno dos valores 0% e 120%, com poucas escolas obtendo valores intermediários. Este fato pode indicar que para algumas escolas as metas são muito difíceis de serem alcançadas, enquanto que para outras, as metas são facilmente atingíveis, uma característica ruim do sistema. Ou seja, apesar de na média, o sistema não ser tão rígido, pode haver uma assimetria de incentivos para diferentes escolas.

4. METODOLOGIA

Para averiguar o impacto do programa, deveríamos comparar o desempenho das escolas em duas situações, com e sem a presença do bônus. No entanto, é impossível observar essas duas situações simultaneamente. Segundo [Abadie, Drukker, Herr e Imbens \(2004\)](#), se a exposição à política for aleatória, pode-se comparar dois grupos: um que recebeu o incentivo (grupo de tratamento) e outro que não (grupo controle). Depois, bastaria calcular a diferença do desempenho nas duas situações para avaliar o impacto do bônus. Entretanto, em nosso estudo, a escolha não é aleatória, uma vez que apenas escolas da rede estadual de São Paulo receberam essa política de incentivos em 2008.

Tabela 2. Porcentagem de Alcance da Meta e Fração de Escolas Não Contempladas.

Ano	% de Alcance da Meta Média		% de Escolas Não Contempladas	
	5º ano	9º ano	5º ano	9º ano
2008	104,9	105,4	49,0	44,2
2009	110,0	100,1	14,0	16,9
2010	108,8	94,3	28,9	44,6
2011	98,6	78,5	22,7	29,6
2012	89,6	66,7	31,2	33,9

* Média calculada excluindo-se escolas que alcançaram 0% da meta.

Fonte: Elaboração própria a partir de dados administrativos da Secretaria de Educação de São Paulo.

Assim, iremos combinar duas técnicas para estimar o efeito do programa de bônus paulista. Inicialmente, será utilizada a técnica de propensity score matching para encontrar para cada escola estadual paulista, uma outra que não pertencesse a essa rede, mas que fosse a mais parecida possível em uma série de características observáveis.

No entanto, mesmo após esse processo, ainda podem existir características não observáveis diferentes (e que afetem a evolução da proficiência escolar) nas escolas da rede estadual paulista e nas demais. Se essas características forem fixas ao longo do tempo, podemos eliminá-las através do método de Diferenças em Diferenças. Além disso, efeitos comuns entre os grupos de tratamento e controle, não causados pelo tratamento, também são eliminados nesse método, como efeitos macro, por exemplo. A combinação dessas estratégias irá permitir estimar o efeito tratamento sobre os tratados, ou seja, o impacto da política de bonificação sobre a proficiência escolar dos alunos do ensino fundamental da rede estadual paulista.

Para que essa combinação de estratégias seja válida na obtenção do efeito causal do tratamento, são necessárias duas hipóteses: suporte comum e tendência comum paralela pré-tratamento. A hipótese de suporte comum garante que nenhuma combinação das covariadas utilizadas no pareamento define exatamente a probabilidade de uma escola pertencer à rede estadual paulista (ou não). A hipótese de tendência comum paralela pré-tratamento exige que os grupos de tratamento e controle tenham evolução paralela nas variáveis de resultado. Em nosso caso, é necessário que a proficiência média das escolas estaduais paulistas tenha trajetória pré-tratamento paralela àquela das escolas do grupo de controle.

Mais formalmente, seja $Y_{it_1}^0$ a proficiência média para a escola i no período t_1 (posterior à adoção da política de bonificação) na ausência da política (sobrescrito “0”). O efeito da política de bônus sobre os tratados pode ser representado como

$$DD = E \left[Y_{it_1}^1 - Y_{it_0}^0 \mid T_i = 1, X_i \right] - E \left[Y_{it_1}^1 - Y_{it_0}^0 \mid T_i = 0, X_i \right]. \quad (4)$$

Note que $E \left[Y_{it_1}^1 \mid T_i = 0, X_i \right]$ é não observado e representa qual seria o resultado médio das escolas do controle, após o tratamento, caso fossem estaduais paulistas. Para que esse contrafactual seja válido para estimação do efeito causal, é necessário que a evolução temporal dos grupos de tratamento e controle, na ausência do tratamento, seja exatamente a mesma. Ou seja,

$$E \left[Y_{it_1}^0 - Y_{it_0}^0 \mid T_i = 1, X_i \right] = E \left[Y_{it_1}^0 - Y_{it_0}^0 \mid T_i = 0, X_i \right]. \quad (5)$$

Como temos múltiplos períodos de tempo, um antes (2007) e dois depois do tratamento (2009 e 2011), a especificação da regressão que permitirá estimar o efeito do tratamento sobre os tratados é dada pela seguinte expressão, adaptada a partir de [Galiani, Gertler e Schargrodsky \(2005\)](#):

$$\text{NotaMédia}_{it} = \beta_0 + \beta_1 EE_i^{\text{SP}} + \beta_2 I_{i2009} + \beta_3 I_{i2011} + \beta_4 d_{2009} + \beta_5 d_{2011} + \varepsilon_{it}, \quad (6)$$

em que NotaMédia_{it} refere-se à proficiência média em cada escola i no tempo t ; $EE_i^{\text{SP}} = 1$, se a escola i for estadual paulista, e 0, caso contrário; I_{i2009} é uma variável que assume valor 1 para as escolas estaduais paulistas apenas em 2009; I_{i2011} assume valor 1 para escolas estaduais paulista apenas em 2011; d_{2009} e d_{2011} são dummies de tempo para os anos de 2009 e 2011, respectivamente.

O intercepto β_0 é a proficiência média das escolas do grupo de controle em 2007. As dummies de tempo capturam mudanças temporais que afetam igualmente o tratamento e o controle. Os parâmetros β_2 e β_3 capturam o efeito causal da política de bonificação sobre o desempenho dos estudantes da rede estadual paulista nos anos de 2009 e 2011, respectivamente.

O método de Diferenças em Diferenças pressupõe que a nota média das escolas estaduais paulistas tenha a mesma trajetória (Hipótese de tendência comum) da nota do grupo de controle. Todavia, ainda pode haver fatores (não relacionados ao programa de bonificação) que afetem as notas dos grupos de



forma desigual. Para eliminar essas possíveis diferenças, podemos combinar mais de um grupo de controle por meio das diferenças triplas (DDD):

$$\begin{aligned} \text{NotaMédia}_{it} = & \beta_0 + \beta_1 \text{SP}_i + \beta_2 \text{EE}_i + \beta_3 d_{2009} + \beta_4 d_{2011} + \beta_5 \text{SP}_i d_{2009} + \beta_6 \text{SP}_i d_{2011} + \beta_7 \text{EE}_i d_{2009} \\ & + \beta_8 \text{EE}_i d_{2011} + \beta_9 \text{EE}_i \text{SP}_i + \beta_{10} \text{EE}_i \text{SP}_i d_{2009} + \beta_{11} \text{EE}_i \text{SP}_i d_{2011} + \varepsilon_{it}, \end{aligned} \quad (7)$$

em que $\text{SP}_i = 1$, se a escola pertence ao estado de São Paulo e 0, caso contrário; $\text{EE}_i = 1$, se a escola é estadual. Neste caso, os coeficientes que medem o impacto do programa em 2009 e 2011 são, respectivamente, β_{10} e β_{11} , e podem ser escritos como

$$\begin{aligned} \beta_{10} = & \left[\left(\text{Nota}_{2009}^{\text{E,SP}} - \text{Nota}_{2007}^{\text{E,SP}} \right) - \left(\text{Nota}_{2009}^{\text{M,SP}} - \text{Nota}_{2007}^{\text{M,SP}} \right) \right] \\ & - \left[\left(\text{Nota}_{2009}^{\text{E,F}} - \text{Nota}_{2007}^{\text{E,F}} \right) - \left(\text{Nota}_{2009}^{\text{M,F}} - \text{Nota}_{2007}^{\text{M,F}} \right) \right] \end{aligned} \quad (8)$$

$$\begin{aligned} \beta_{11} = & \left[\left(\text{Nota}_{2011}^{\text{E,SP}} - \text{Nota}_{2007}^{\text{E,SP}} \right) - \left(\text{Nota}_{2011}^{\text{M,SP}} - \text{Nota}_{2007}^{\text{M,SP}} \right) \right] \\ & - \left[\left(\text{Nota}_{2011}^{\text{E,F}} - \text{Nota}_{2007}^{\text{E,F}} \right) - \left(\text{Nota}_{2011}^{\text{M,F}} - \text{Nota}_{2007}^{\text{M,F}} \right) \right] \end{aligned} \quad (9)$$

em que E = Estadual, M = Municipal, SP = Estado de São Paulo e F = Fora do Estado de São Paulo.

5. DADOS

Para avaliarmos o efeito do programa de bonificação sobre a proficiência dos alunos utilizaremos os dados da Prova Brasil. Embora a bonificação esteja atrelada às notas do SARESP, que é a avaliação externa aplicada pela Secretaria de Educação do estado de São Paulo, esses dados limitam nossa avaliação e a aplicação de metodologias de avaliação de políticas públicas. Como o exame é realizado apenas em São Paulo, ela impossibilita que alunos e escolas de outros estados possam servir como grupo de controle.

Assim, será utilizada como base de dados a Prova Brasil. Nela, são realizados testes de Matemática (com foco na resolução de problemas) e de Língua Portuguesa (com foco em leitura) para o quinto e nono anos do ensino fundamental e terceiro ano do ensino médio. Este exame tem características muito semelhantes ao SARESP: usam como base a TRI (Teoria de resposta ao item) e avaliam os anos finais de cada ciclo em pelo menos duas disciplinas básicas (Matemática e Português).

Segundo o [Governo do Estado de São Paulo/SEE \(2009a\)](#), em 2007 foram introduzidas mudanças importantes no SARESP, num esforço para tornar o exame mais adequado tecnicamente às características de um sistema de avaliação em larga escala que permitisse acompanhar a evolução da qualidade do sistema estadual de ensino ao longo dos anos. Além de o SARESP ter adequado as habilidades avaliadas aos padrões do Saeb, a escala de proficiência passou a ser igual ao da prova nacional. Assim, a partir de 2007, os resultados obtidos pelos alunos paulistas nos dois exames ao longo dos anos tornaram-se passíveis de comparação. Dessa forma, a semelhança das avaliações nos leva a acreditar que o efeito de um programa que pretende melhorar o ensino deve ter impacto positivo nas notas dos dois exames.

Além da Prova Brasil nos proporcionar comparações com escolas de outros estados, ela evita alguns tipos de *gaming*⁵. Como este exame não está diretamente atrelado ao pagamento de bônus, não há incentivos para seleção de alunos, ajuda na prova ou outros tipos de fraudes que podem ocorrer no dia

⁵Pode-se definir *gaming* como comportamentos estratégicos intencionais, voltados apenas a aumentar a nota média das provas, sem necessariamente melhorar o aprendizado dos alunos de modo consistente.

da aplicação. Mesmo que não seja possível evitar todos os efeitos indesejados, o uso do exame nacional ajuda a minimizar alguns deles.

Para as estimações, utilizaremos as notas médias das escolas da Prova Brasil de 2007, 2009 e 2011 de Língua Portuguesa e Matemática para o quinto ano e nono ano do ensino fundamental. Como a Prova Brasil não é censitária no 3º ano do ensino médio, excluimos essa série da análise. Além das provas, os alunos, professores e diretores respondem a questionários socioeconômicos que fornecem as covariadas necessárias para a estimação do *Propensity Score*.

A [Tabela 3](#) possui uma breve análise descritiva das variáveis da Prova Brasil de 2007 utilizadas no processo de *Propensity Score Matching*. As diferenças entre as características dos alunos das escolas estaduais paulistas e os alunos de escolas municipais de SP, bem como do resto do Brasil não são muito grandes, tanto no quinto como no nono ano. Já em relação ao salário dos professores, as escolas estaduais de São Paulo mostraram indicadores melhores do que a média do País e do que a média das redes municipais do estado de São Paulo. O mesmo é válido para a diferença salarial entre diretores: enquanto 58% dos diretores em escolas do primeiro ciclo do ensino fundamental recebem mais de R\$ 1.500 no país como um todo, nas redes municipais de SP esse número atinge 86% e na rede estadual paulista, 96%. Por outro lado, os docentes da rede estadual paulista possuem menos tempo de experiência dentro de uma mesma escola no 5º ano; já no 9º ano, isso não ocorre. Também não se verificam grandes diferenças em termos de tempo no cargo de diretor da escola entre os três grupos analisados. Um aspecto interessante é que as escolas estaduais paulistas têm um número maior de alunos, em média, do que as demais, tanto no 5º quanto no 9º ano. Por fim, em termos de proficiência, a rede estadual paulista ocupava em 2007, uma posição intermediária entre o resto do Brasil e as redes municipais de São Paulo. Esse resultado ratifica nossa escolha de trabalharmos com diversos grupos de controles alternativos para assegurar a robustez dos resultados.

6. RESULTADOS

A estimação do escore de propensão, ou seja, da probabilidade de ser tratado, foi realizada por meio do modelo Probit. No contexto deste trabalho, foi estimada a probabilidade de ser uma escola estadual paulista. Nesse conjunto de resultados iniciais, consideramos que qualquer escola do Brasil possa ser, potencialmente, utilizada como controle para as escolas estaduais paulistas.

A vasta disponibilidade de dados da Prova Brasil 2007 nos permite, a princípio, utilizar um número muito grande de variáveis explicativas. Quanto mais covariadas relevantes conseguirmos utilizar, mais o grupo de controle tenderá a ser parecido com o grupo de tratamento; por outro lado, aumenta a dificuldade de encontrar um controle exatamente igual ao tratado. Num primeiro exercício, ao utilizar um conjunto muito grande de variáveis explicativas, poucos indivíduos foram pareados, levando a uma pequena região de suporte comum. Desta forma, optou-se por um modelo com um número menor de covariadas. A [Tabela A-1 \(Apêndice A, página 234\)](#) apresenta os resultados da estimação das especificações que apresentaram uma boa qualidade do matching, com covariadas balanceadas e maior região de suporte comum. Cabe destacar que foram utilizados como regressores do Probit a proficiência média em 2007 e a variação da proficiência entre 2005 e 2007, no intuito de garantir a validade da hipótese de tendência paralela pré-tratamento na amostra pareada.

Após a estimação do escore de propensão foi feito o pareamento através do método de um vizinho mais próximo e impondo a condição de suporte comum. A [Tabela 4](#) mostra que o número de escolas tratadas dentro do suporte comum é bastante grande, variando entre 93% a 96% das escolas, dependendo da série e da disciplina avaliada.

A boa aderência do pareamento também pode ser notada ao observar a distribuição do escore de propensão ([Figura 1](#)) após o pareamento. Outro procedimento importante nesse tipo de metodologia é a checagem das condições de balanceamento. As tabelas [A-2 a A-5 no Apêndice A](#) mostram a média das variáveis nos grupos de tratamento e controle. Depois do pareamento, para a maioria das covariadas não foi possível rejeitar a hipótese nula de igualdade das médias e, portanto, tem-se um pareamento



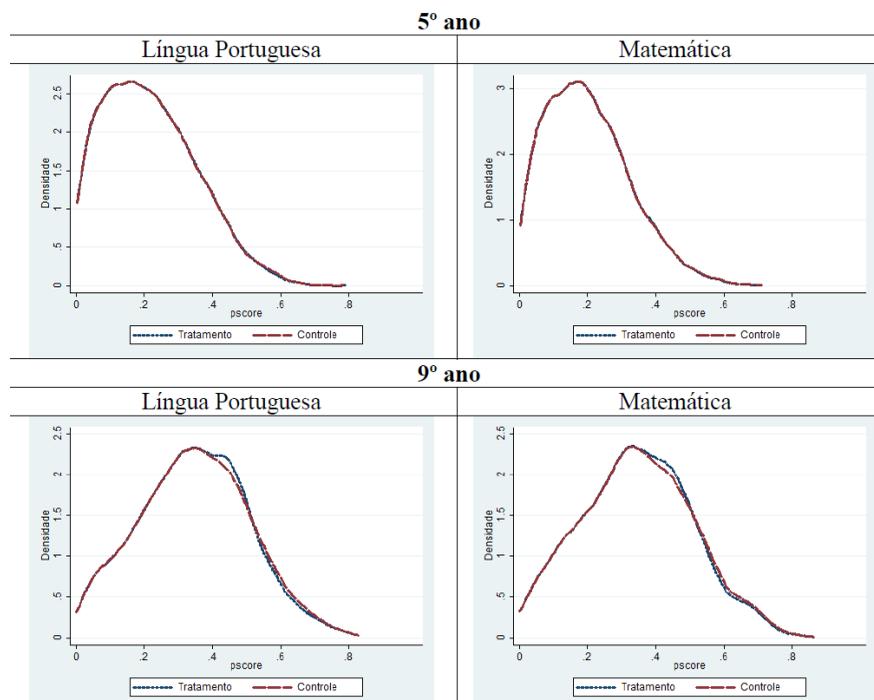
Tabela 3. Médias das variáveis utilizadas no modelo probit para Brasil, Rede Estadual de SP e Redes Municipais de São Paulo em 2007.

	Quinto ano			Nono ano		
	Brasil	SP (Estadual)	SP (Municipal)	Brasil	SP (Estadual)	SP (Municipal)
<i>Proporção de alunos com a característica dentro da escola</i>						
Mãe completou a 8ª série	12,5	14,1	13,7	15,5	18,4	18,1
Possuem de 1 a 20 livros em casa	52,1	52,4	53,0	59,3	60,3	60,3
Possuem de 21 a 100 livros em casa	15,7	15,5	15,3	19,0	18,3	17,9
Possuem mais de 100 livros em casa	9,1	8,9	8,2	6,3	5,5	5,0
Fizeram o Maternal	33,2	34,0	35,9	41,5	42,2	39,5
Trabalham fora de casa	13,8	13,1	11,6	20,9	18,3	17,7
Mora com o pai e com a mãe	53,7	62,0	63,5	59,7	61,8	62,1
Mora com o pai ou com a mãe	33,4	32,2	30,8	29,5	32,1	31,7
<i>Proporção de professores com a característica dentro da escola</i>						
Salário até R\$ 1.300,00	30,9	8,1	12,3	25,5	6,6	9,3
Leciona na escola entre 1 e 2 anos	14,0	20,4	15,8	15,2	18,8	17,0
Leciona na escola entre 3 e 5 anos	21,7	11,7	23,9	21,2	20,4	26,5
<i>Variáveis de Diretores/Escola</i>						
Salário maior que R\$ 1.500,00	57,1	95,7	86,2	69,0	97,2	89,9
Diretor da escola há até 2 anos	33,1	34,9	32,4	35,5	34,4	32,7
Diretor da escola entre 2 e 4 anos	35,3	32,5	35,4	32,4	31,3	32,8
Diretor da escola entre 5 e 10 anos	22,8	24,1	25,2	23,1	24,0	25,5
Média de nº de alunos da escola	483	668	525	730	969	826
<i>Proficiência Escolar</i>						
Média de Matemática em 2007	188,71	193,6	203,44	238,86	242,03	246,44
Média de Português em 2007	171,29	176,48	184,03	227,16	231,13	233,87

Fonte: Elaboração própria com base nos dados do Inep.

Tabela 4. Distribuição de Escolas por região de suporte comum após pareamento.

		Suporte Comum		
		Fora do suporte	Dentro do suporte	Total
<i>Quinto ano</i>				
Português	Controles	0	21.780	21.780
	Tratados	49	1.383	1.432
Matemática	Controles	0	21.774	21.774
	Tratados	8	1.423	1.431
<i>Nono ano</i>				
Português	Controles	0	14.628	14.628
	Tratados	201	2.734	2.935
Matemática	Controles	0	14.632	14.632
	Tratados	181	2.749	2.930

Figura 1. Densidade Kernel do escore de propensão após pareamento – 5º e 9º anos – Português e Matemática.

com um bom balanceamento dessas variáveis.

Por fim, foi checado o que ocorreu com o Pseudo R^2 do modelo Probit após o pareamento. Ao se reestimar o Probit utilizando apenas informações das escolas tratadas e de controle selecionadas, o valor do Pseudo R^2 deve diminuir se o pareamento foi bem sucedido. É exatamente o que ocorre em nosso caso, como mostrado na [Tabela 5](#), já que os grupos passam a ser semelhantes nessas covariadas. O poder explicativo das covariadas cai, ou seja, elas não conseguem mais explicar as diferenças existentes entre grupo de controle e tratamento e a probabilidade de participação.

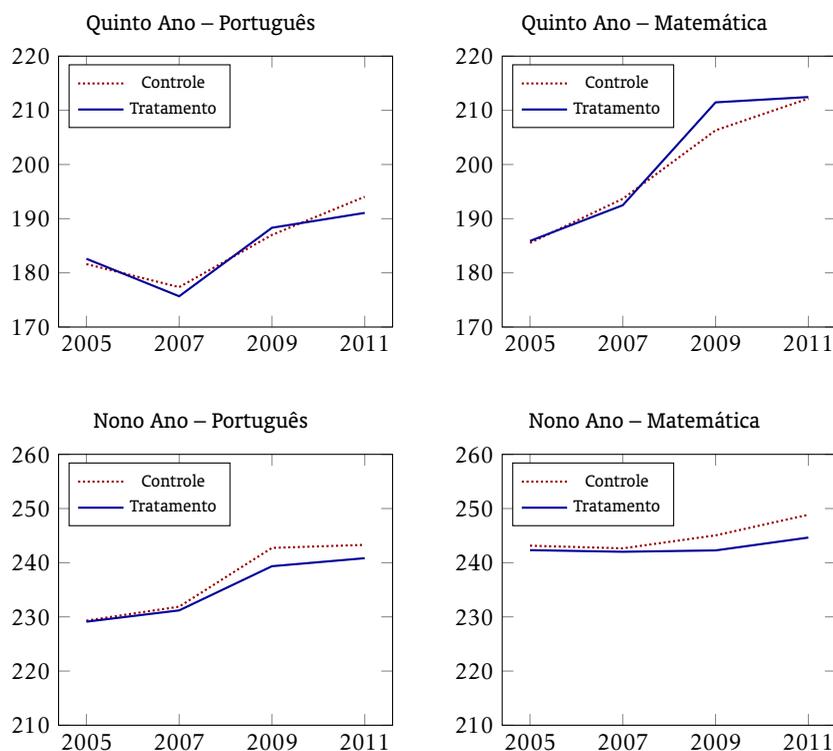
A [Figura 2](#) apresenta a evolução das notas de português e matemática, para o 5º e 9º ano, na amostra de escolas de tratamento e controle após o pareamento. Conforme a [Tabela 4](#), mais de 93% das escolas de tratamento estão sendo utilizadas para a análise, percentual que chega a 99% no caso de matemática no quinto ano. Ou seja, essa evolução reflete de modo fidedigno o que ocorreu na rede estadual paulista nos últimos anos. O primeiro aspecto a ser destacado nesse gráfico é que a evolução das notas das escolas do grupo de controle é bastante semelhante daquela observada para as escolas de tratamento entre 2005 e 2007. Assim, a hipótese do método de diferenças em diferenças de tendência comum pré-tratamento, fundamental para a inferência causal nesse método, é respeitada no nosso caso. Infelizmente,

Tabela 5. Pseudo R^2 .

	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
Antes do pareamento	0,293	0,268	0,267	0,263
Depois do pareamento	0,007	0,001	0,002	0,003



Figura 2. Evolução das Notas Médias de Português e Matemática na Prova Brasil – Quinto e Nono ano – 2005 e 2011.



Obs: Estão representadas na figura a proficiência média para as escolas do grupo de tratamento (estaduais paulistas) e do grupo de controle após o pareamento realizado via *propensity score matching*.
 Fonte: Elaboração própria com base nos dados da Prova Brasil.

não dispomos de informações anteriores à 2005 para assegurar que tal comportamento ocorria antes dessa data.

O segundo ponto a ser destacado é que entre 2007 e 2011, no 5º ano, houve crescimento das notas de português (12 pontos) e matemática (20 pontos), sendo essa evolução bastante semelhante nas escolas de tratamento e de controle. No caso de matemática, a rede estadual experimentou um avanço mais forte até 2009 e, entre 2009 e 2011, as escolas de controle eliminaram essa diferença. Já no 9º ano, as escolas do grupo de controle tiveram desempenho ligeiramente melhor que as de controle tanto em 2009 quanto em 2011. Um dado preocupante adicional no caso de matemática no 9º ano é que praticamente não houve evolução da proficiência média das escolas estaduais entre 2005 e 2011. Desse modo, as evidências da [Figura 1](#) já sinalizam que, provavelmente, a política de bonificação não conseguiu mudar radicalmente a tendência do aprendizado das escolas estaduais paulistas após sua implantação.

No entanto, para averiguarmos se essa evidência gráfica se traduz em resultados estatísticos, foi aplicado o estimador de Diferenças em Diferenças na amostra pareada para eliminar o efeito de características não observadas (fixas no tempo). Os resultados mostram indícios de impactos positivos apenas para o quinto ano do ensino fundamental e limitado ao ano de 2009 ([Tabela 6](#)). No exame de Matemática, o efeito estimado foi de cerca de 6,3 pontos na prova, o que representa 0,42 desvio padrão na escala SAEB, o que é um efeito considerável. Já para o exame de Português, o efeito do bônus mostrou-se menor, em cerca de 2,6 pontos (0,14 desvio padrão). No entanto, esse efeito desaparece quando se

consideraram os dados de 2011. Ou seja, para o quinto ano, há efeito de curto prazo, mas o decaimento é bastante pronunciado.

Com relação ao nono ano, os coeficientes são negativos, indicando que a política de bonificação pode estar contribuindo para diminuir a proficiência dos alunos.

Outro dado que chama atenção na Tabela 6 é que houve evolução da proficiência em ambos os grupos (tratamento e controle) no período, de mais de 10 pontos na escala SAEB entre 2007 e 2011, o que pode ser verificado nos coeficientes das variáveis d_{2009} e d_{2011} . Por sua vez, é alarmante é a tímida evolução da proficiência em matemática no 9º ano.

7. ANÁLISE DE ROBUSTEZ

Uma das críticas que pode ser feita aos resultados anteriormente encontrados é que eles podem ser sensíveis à escolha do grupo de controle, ou seja, se mudássemos o grupo de controle, os resultados poderiam ser diferentes. Implicitamente, essa crítica aponta que o grupo de controle escolhido através do pareamento pode não ser um bom contrafactual para o grupo de escolas estaduais paulistas. Para checar a robustez dos resultados encontrados, nessa seção serão utilizadas técnicas e grupos de controle alternativos para checarmos se os resultados encontrados são robustos a alterações nessas especificações. Em todos os casos (exceto Diferenças Triplas e Diferenças em Diferenças por municípios), foi realizado o pareamento das escolas de tratamento e de controle pré intervenção, por meio do *propensity score matching*.⁶ Também foram construídas figuras análogas à Figura 2, para assegurar a validade da hipótese de tendência comum pré-intervenção, o que foi constatado em praticamente todos os casos.⁷ Relacionamos abaixo as estratégias para assegurar a robustez dos resultados.

Tabela 6. Resultado do modelo de Diferenças em Diferenças.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
EE ^{SP}	-0,13 (0,59)	-0,34 (0,63)	-0,74* (0,39)	-0,33 (0,41)
d_{2009}	10,19*** (0,65)	13,60*** (0,76)	11,11*** (0,44)	2,74*** (0,47)
d_{2011}	17,09*** (0,67)	19,61*** (0,77)	11,75*** (0,45)	6,84*** (0,48)
I_{2009}	2,55*** (0,86)	6,29*** (0,97)	-2,36*** (0,58)	-2,11*** (0,6)
I_{2011}	-1,82** (0,87)	1,15 (0,98)	-1,67*** (0,58)	-3,71*** (0,62)
Constante	177,83*** (0,45)	194,90*** (0,49)	232,22*** (0,3)	242,97*** (0,32)
Observações	8.298	8.538	16.404	16.494
R^2	0,1501	0,1834	0,0968	0,0226

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

⁶As variáveis utilizadas nos pareamentos foram as mesmas utilizadas nos pareamentos da seção 6. O resultado da estimação dos modelos Probit, bem como a análise de qualidade desses pareamentos podem ser obtidas por solicitação aos autores.

⁷As figuras A-2 à A-6 do Apêndice A apresentam a evolução das notas dos grupos de tratamento e controle nesses casos.



Diferenças em diferenças ponderado pelo escore de propensão. Seguindo Hirano, Imbens e Ridder (2003), a equação (6) foi reestimada, ponderando-se as observações pelo escore de propensão estimado. Ou seja, a estimação do *propensity score* e o pareamento foi exatamente o mesmo realizado na seção 6. Mudou-se apenas a ponderação do estimador de diferenças em diferenças.

Diferenças em diferenças por município. O modelo do programa permite que dentro de um mesmo município existam escolas que foram contempladas pelo programa (rede estadual) e escolas que não tiveram o incentivo da bonificação (outras redes). Escolas de um mesmo município estão expostas aos mesmos fatores econômicos e geográficos, bem como ao mesmo ambiente macroeconômico. É razoável supor que esses aspectos afetem de forma semelhante estudantes de redes de ensino distintas dentro de um mesmo município. Assim, a comparação entre escolas da mesma localidade pode mostrar melhor o efeito do programa, já que se estaria controlando por efeitos locais que afetam a proficiência escolar nas duas redes. Devido à pouca disponibilidade de dados para escolas particulares e federais, o grupo de comparação será formado apenas por escolas municipais.

Escolhido o contrafactual, lançaremos mão, novamente, da hipótese de tendência comum e, ao considerar fixos os efeitos dentro de cada município, usaremos então o método de Diferenças em Diferenças. A equação (6) foi acrescida de variáveis dummy para cada município e pode ser lida pela seguinte expressão:

$$\text{NotaMédia}_{it} = \beta_0 + \beta_1 \text{EE}_i^{\text{SP}} + \beta_2 I_{i2009} + \beta_3 I_{i2011} + \beta_4 d_{2009} + \beta_5 d_{2011} + \text{Município}_1 + \text{Município}_2 + \dots + \text{Município}_J + \varepsilon_{it}, \quad (10)$$

na qual J representa a quantidade de municípios que serão considerados. O estado de São Paulo contém 645 municípios, mas para esta estratégia foram utilizadas apenas cidades paulistas com pelo menos três escolas de cada rede de ensino. Ao todo, foram utilizados dados de 83 municípios para a estimação do quinto ano e 42 municípios para o nono ano.

Escolas estaduais paulistas × Escolas municipais paulistas. Nesse caso, no processo de pareamento, foram usados como potenciais controles apenas escolas municipais localizadas no estado de São Paulo, mas sem impor a restrição de que cada par de escolas tratamento-controle tivesse que ser, necessariamente, do mesmo município como no caso anterior. Uma vez feito o pareamento, foi estimada a equação (6) na amostra pareada.

Escolas estaduais paulistas × Escolas estaduais do resto do Brasil. Foram utilizadas como possíveis contrafactuais apenas escolas pertencentes a redes estaduais das outras unidades de federação e estimada a equação (6) na amostra pareada.

Escolas estaduais paulistas × Escolas estaduais de MG, RJ e ES. Também foi utilizado como possível contrafactual o grupo formado por escolas estaduais da região Sudeste (exceto do estado de São Paulo) já que é razoável admitir que escolas de estados vizinhos tenham maior probabilidade de serem pareadas em fatores não observáveis que escolas muito distantes.

Escolas estaduais paulistas × Escolas estaduais de RJ e ES. Esse grupo de controle é semelhante ao anterior, excluindo-se Minas Gerais pelo fato de as notas de 5º ano em Minas Gerais terem comportamento distinto do observado para os demais estados do Sudeste.

Escolas estaduais paulistas × Escolas de redes que não introduziram política de bonificação após 2008. Uma crítica que pode ser feita aos grupos de controle acima é que essas escolas podem pertencer a redes estaduais que também adotaram políticas de bonificação aos docentes após a entrada em vigor

do programa paulista. Desse modo, após levantamento das políticas de bônus adotadas nos últimos anos nas redes estaduais brasileiras, excluimos escolas das redes estaduais dos seguintes estados: RJ, MG, ES, PE, GO e CE. Claramente, uma limitação dessa análise é que não temos como saber que redes municipais adotaram políticas de bônus semelhantes àquela do estado de São Paulo. Assim, é possível que haja algumas escolas no grupo de controle que também estejam suscetíveis a políticas de bonificação.

Efeito Placebo. Outro exercício de robustez realizado foi estimar modelo de diferenças em diferenças da equação abaixo com dados de 2005 e 2007, considerando todas as escolas do Brasil como potenciais controles:

$$\text{NotaMédia}_{it} = \beta_0 + \beta_1 EE_i^{\text{SP}} + \beta_2 I_{i2007} + \beta_3 d_{2007} + \varepsilon_{it}. \quad (11)$$

Estimamos, nesse caso, um efeito placebo, isto é, qual teria sido o resultado de uma política de bônus adotada entre 2005 e 2007. Como na verdade a política de bonificação data de 2008, espera-se que o coeficiente β_2 seja não significativo na (11).

Diferenças triplas (DDD). O método de Diferenças em Diferenças exige que as trajetórias das notas das escolas estaduais paulistas sejam semelhantes às dos grupos de controle. Todavia, ainda pode haver fatores (não relacionados ao programa de bonificação) que afetem as notas dos grupos de forma desigual como, por exemplo, as diferenças inerentes entre as administrações estaduais e municipais. Para eliminar essas possíveis diferenças, podemos combinar mais de um grupo de controle por meio das diferenças triplas (DDD). Um aspecto importante a ser considerado é que, nesse caso de Diferenças Triplas, não foi realizado o processo de pareamento das escolas de tratamento e controle antes da estimação do modelo, o que pode ocasionar alguma diferença em relação às análises anteriormente citadas. Foram realizados três exercícios de DDD, com os seguintes grupos:

- Escolas Estaduais SP vs Escolas Municipais SP × Escolas Estaduais Brasil vs Escolas Municipais Brasil;
- Escolas Estaduais SP vs Escolas Municipal SP × Escolas Estaduais Sudeste vs Escolas Municipais Sudeste;
- Escolas Estaduais SP vs Escolas Municipais SP × Escolas Estaduais RJ + ES vs Escolas Municipais RJ + ES.

7.1. Diferenças em Diferenças e Efeito Placebo

A [Tabela 7](#) apresenta os coeficientes de impacto causal para as análises de robustez mencionadas anteriormente e que envolvem o uso do método de diferenças em diferenças.⁸

Os resultados mais robustos foram encontrados para o quinto ano em matemática em 2009. Os valores de efeito causal em 2009 variaram entre 2,88 e 7,71 pontos na escala SAEB, dependendo do grupo de controle considerado. No entanto, os resultados para 2011 dessa disciplina e série indicam desde valores não significativos até o valor de 6,16 (significativo), no caso de diferenças em diferenças dentro dos municípios paulistas. Cabe destacar que, nesse caso, o efeito está limitado a um subconjunto de municípios que possuem pelo menos 3 escolas de cada rede de ensino. Assim, se por um lado, as escolas estão sendo comparadas com outras da mesma cidade, por outro a validade externa do resultado é menor. De qualquer modo, em todos os casos há uma diminuição (ou desaparecimento em quatro especificações) do efeito do programa entre 2009 e 2011, corroborando o resultado da [Tabela 6](#).

Para português no quinto ano em 2009, apenas duas especificações não mostraram efeito positivo e significativo: quando o grupo de controle foi restrito às escolas municipais paulistas e quando se restringiu às escolas estaduais do sudeste. Nos outros cinco casos, o efeito foi positivo e variou entre 1,58

⁸Os resultados completos de todas as estimações encontram-se nas tabelas A-6 a A-12 do [Apêndice A](#).

**Tabela 7.** Resultado da análise de robustez – Diferenças em Diferenças.

	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
D-D com reponderação				
I_{2009}	2,36***	6,01***	-2,44***	-2,09***
I_{2011}	-1,41	1,54*	-1,67***	-3,74***
Diferenças em diferenças por município				
I_{2009}	4,02***	7,25***	-1,14	-0,19
I_{2011}	2,34***	6,16***	-2,51***	-2,59***
Grupo de controle: Escolas municipais paulistas				
I_{2009}	1,26	4,00***	-0,35	0,32
I_{2011}	-1,49	1,72	-3,86***	-3,36***
Grupo de controle: Escolas estaduais do restante do Brasil				
I_{2009}	2,59***	7,71***	-2,21***	-2,11***
I_{2011}	-2,42**	2,16**	-0,30	-2,40***
Grupo de controle: Escolas estaduais do restante do sudeste				
I_{2009}	-0,16	2,88***	-2,61***	-3,14***
I_{2011}	-2,69***	-0,19	-3,64***	-6,09***
Grupo de controle: Escolas estaduais do RJ e ES				
I_{2009}	1,58*	4,85***	-1,42	-1,13
I_{2011}	-1,42	1,46	-1,26	-3,38***
Grupo de controle: Escolas do restante do Brasil, exceto das redes estaduais de MG, RJ, CE, PE, ES e GO				
I_{2009}	1,85**	5,92***	-2,80***	-2,28***
I_{2011}	-2,54***	0,41	-2,26***	-3,59***
Efeito Placebo				
I_{2007}	-1,56*	-1,06	-0,51	0,03

Notas: Estimação com erro padrão robusto a heterocedasticidade.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

e 4,02 pontos na escala SAEB. Assim, nesse caso, não corroboramos o resultado negativo e significativo da Tabela 6.

Para português no nono ano em 2009, as estimativas ou são não significativas ou significativas com sinal negativo, sendo que o valor encontrado na Tabela 6 (-2,36) encontra-se dentro desse intervalo. Como há três casos não significativos, o resultado de que a política de bonificação teria efeito *negativo* sobre a proficiência escolar não é robusto para o ano de 2009. No entanto, ao analisarmos o efeito em 2011, em apenas dois casos o efeito é não significativo; nos demais é negativo e significativo. Assim, as evidências nos levam a concluir que o efeito do bônus em 2011 foi, na melhor das hipóteses nulo, mas é provável que tenha sido negativo em até 3,86 pontos na escala SAEB. Para matemática no nono ano em 2009, há alguns resultados negativos, mas outros não significativos. No entanto, para 2011, todas as especificações indicam efeito *negativo* entre 2,4 e 6,9 pontos na escala SAEB, corroborando os resultados da Tabela 6.

Por fim, a estimativa do efeito placebo corroboram parcialmente a nossa hipótese de identificação, já que para Português, no quinto ano, há um efeito negativo e significativo de cerca de 1,5 pontos na escala SAEB. Já para os demais casos, os coeficientes são não significativos, conforme esperado.

Portanto, os resultados da análise de robustez usando diferenças em diferenças indicam que o efeito mais forte da política de bônus ocorreu para o quinto ano em matemática em 2009, um efeito menor

ocorreu em português, mas em ambos os casos, houve um decaimento importante do efeito em 2011. Para o nono ano, os resultados apontam para uma situação em que ou pode estar ocorrendo efeito negativo ou a ausência de efeito da política em 2009. Porém, em 2011, os resultados indicam para efeito negativo de 3 pontos na escala SAEB (média das estimativas da Tabela 7).

7.2. Diferenças triplas (DDD)

Os resultados de diferenças triplas apontam um quadro um pouco mais positivo para o programa, com valores positivos, significativos e maiores do que os anteriormente encontrados, tanto para matemática como para português em 2009 e 2011 no 5º ano. Já para o 9º ano, os resultados confirmam os anteriormente obtidos de ausência de efeito positivo do programa, tanto em 2009 como em 2011.

7.3. Bônus e Indicadores Escolares

Mas o que pode explicar os resultados obtidos e as diferenças entre 5º e 9º anos? Apesar de não termos medidas diretas do comportamento dos docentes, pode ser que a política de bonificação tenha induzido mudanças nas escolas (perfil dos alunos, diretor e docentes, clima escolar, práticas didáticas, etc.), que podem estar mediando o impacto do bônus sobre as notas. Assim, construímos o mesmo conjunto de indicadores de Tavares (2014) e utilizamos cada um desses indicadores como variável dependente na equação (6), no lugar da proficiência escolar dos alunos. Assim, conseguimos saber se esses indicadores mudaram mais nas escolas estaduais paulistas do que nas escolas do grupo de controle entre 2007 e 2011.

Para o quinto ano em 2009, houve melhoria do clima escolar, maior liderança do diretor e maior participação dos pais nas escolas e maior alocação de alunos em turmas homogêneas (por idade e rendimento acadêmico). Essas variáveis podem explicar o porquê houve um efeito positivo do bônus sobre as notas, tanto de português como de matemática em 2009. Por outro lado, entre 2007 e 2009, relativamente às escolas do grupo de controle, as estaduais paulistas tiveram uma piora do clima escolar, da

Tabela 8. Resultado da análise de robustez – Diferenças Triplas.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
<i>Estadual SP vs Municipal SP × Estadual Br vs Municipal Br</i>				
EE-SP-2009	4,06*** (0,70)	5,43*** (0,80)	-0,33 (0,83)	0,63 (0,89)
EE-SP-2011	2,37*** (0,72)	4,86*** (0,82)	-0,78 (0,82)	-0,61 (0,90)
<i>Estadual SP vs Municipal SP × Estadual SE vs Municipal SE</i>				
EE-SP-2009	5,78*** (0,89)	5,71*** (1,02)	-2,32** (1,01)	-1,37 (1,11)
EE-SP-2011	5,74*** (0,90)	8,83*** (1,02)	-0,01 (1,00)	0,45 (1,13)
<i>Estadual SP vs Municipal SP × Estadual RJ + ES vs Municipal RJ + ES</i>				
EE-SP-2009	8,12*** (1,21)	9,45*** (1,34)	-1,18 (1,25)	-1,87 (1,33)
EE-SP-2011	8,58*** (1,28)	10,82*** (1,41)	4,38*** (1,26)	3,10** (1,40)

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.



formação acadêmica dos docentes e redução dos diretores com pós-graduação na área de gestão, fatores que podem ter atuado para diminuir o impacto da política de bonificação nesse período.

Entre 2009 e 2011, o efeito do programa sobre a proficiência se reduz bastante no quinto ano. Curiosamente, a maioria dos indicadores passa a ser impactado positivamente em 2011, o que deveria contribuir para ampliar o efeito da política. Por exemplo, houve melhoria no clima escolar, na liderança do diretor, no uso do tempo em sala de aula, no cumprimento do currículo, diminuição do número de alunos médio por escola e maior dedicação do professor. Os indicadores na direção contrária foram: piora no índice de composição socioeconômica dos alunos, piora na participação dos pais, menos diretores com pós-graduação em gestão e diminuição do tempo de experiência do diretor.

Portanto, para o quinto ano, não fica claramente estabelecido um padrão dos indicadores que nos permita associá-los diretamente ao processo de melhoria inicial e posterior decaimento da proficiência escolar após a implantação do bônus em 2008.

Para o nono ano, onde os resultados sobre proficiência para ambos os anos foram negativos ou inexistentes, os sinais dos efeitos da política de bônus sobre os indicadores foram semelhantes aos encontrado para o quinto ano em praticamente todos os casos. Assim, não parece que os indicadores observáveis conseguem explicar o mau resultado da política de bonificação no nono ano, assim como não explicavam os resultados mais favoráveis do quinto ano.

8. CONSIDERAÇÕES FINAIS

As reformas educacionais têm como um dos temas centrais programas de responsabilização dos agentes escolares pela melhoria do desempenho dos alunos. Dentro deste contexto, este trabalho procurou investigar se o programa de bonificação de professores implantado em 2008 no estado de São Paulo teve impacto sobre a proficiência dos estudantes.

Boa parte da literatura internacional indica que programas de premiação de docentes levam a uma melhora do desempenho dos alunos, porém, há estudos que não mostram impacto sobre as notas, ou até mesmo, que mostram efeitos negativos, como Fryer (2013).

Para este trabalho, utilizou-se como método principal o Diferenças em Diferenças (DD) para eliminar efeitos fixos não observáveis, combinado com Propensity Score Matching (PSM) aplicado no período anterior ao início da política para encontrar escolas para o grupo de controle que fossem os mais semelhantes possíveis com as escolas estaduais paulistas (grupo de tratamento). Procedimentos de checagem mostraram boa qualidade do pareamento, bem como a validade da hipótese chave do método de diferenças em diferenças de tendência comum prévia ao tratamento na variável de resultado para as escolas de tratamento e de controle.

Foram encontrados indícios de impacto positivo do programa de incentivo de professores apenas nas notas do quinto ano do ensino fundamental em 2009, de 6,3 pontos na escala SAEB em Matemática e de 2,6 pontos em Língua Portuguesa. No entanto, há forte decaimento desses efeitos entre 2009 e 2011. Para o nono ano, encontramos resultados negativos; mas nas análises de robustez para 2009, em alguns casos observou-se ausência de efeito da política. No entanto, para 2011, praticamente todas as estimativas confirmaram um impacto negativo.

Por fim, tentamos avaliar se a adoção do bônus alterou alguns indicadores escolares (como clima escolar, participação dos pais na escola, cumprimento do currículo, uso do tempo em sala de aula, experiência dos professores/diretores, etc.) que poderiam ajudar a entender tanto a diferença de resultados entre o quinto e o nono anos, como o decaimento do efeito para o quinto ano entre 2009 e 2011. No entanto, essa análise foi inconclusiva em apontar claramente algum desses fatores como estando associados a essa trajetória do efeito da política sobre as notas.

Nesse sentido, é possível levantar algumas hipóteses adicionais que poderiam ajudar a explicar as diferenças de impacto sobre notas tanto entre as disciplinas e como nos diferentes anos do ensino fundamental. Com relação à diferença entre Português e Matemática, devemos considerar a diferença de aprendizado nas duas disciplinas. Talvez o esforço do professor dentro de sala de aula seja mais efetivo

Tabela 9. Efeito da política de bonificação sobre indicadores escolares – 5º ano.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
<i>Contínuas</i>				
Índice socioeconômico (+)	0,00	-0,21 **	0,02	-0,22 **
Índice de clima escolar (-)	0,24 *	-0,35 ***	0,30 **	-0,22 **
Expectativa do professor sobre futuro dos alunos (+)	0,03	0,00	0,05 **	0,00
Índice de liderança do diretor (+)	0,43 ***	0,51 ***	0,30 **	0,50 ***
Índice de uso do tempo de aula (+)	0,06	0,23 ***	0,08 *	0,25 ***
Índice de cumprimento curricular	0,11 ***	0,10 ***	0,12 ***	0,09 ***
Formação acadêmica do professor (+)	-0,17 ***	-0,20 ***	-0,18 ***	-0,21 ***
Experiência do professor (+)	-0,02	0,04 *	-0,01	0,07 ***
Número total de matrículas	-40,09	-79,56 ***	-39,87	-75,90 ***
Índice de Participação dos pais na escola (+)	0,27 ***	-0,07 ***	0,25 ***	-0,08 ***
Dedicação do professor (+)	0,00	0,08 ***	-0,02	0,07 **
<i>Discretas</i>				
Formação do diretor (<i>dummy</i> pós-graduação em gestão)	-0,07 ***	-0,06 **	-0,08 ***	-0,06 **
Experiência do diretor (<i>dummy</i> mais de 5 anos)	0,00	-0,13 ***	-0,01	-0,15 ***
Alocação dos alunos em turma homogênea (por idade e rendimento)	0,06 **	0,22 ***	0,04	0,21 ***

Notas: Os valores correspondem aos coeficientes estimados das variáveis I_{i2009} e I_{i2011} da equação (6), onde cada indicador contido na primeira coluna da tabela foi utilizado como variável dependente dessa equação. Para detalhamento e definição dos indicadores, ver Tavares (2014). O sinal (+) indica que aumentos no valor do índice levam a melhoria no indicador; (-) tem a interpretação inversa.

Fonte: Elaboração própria com base nos dados da Prova Brasil de 2007, 2009 e 2011.

para o aprendizado de Matemática do que de Língua Portuguesa, uma vez que, para esta última, a leitura fora de sala de aula é bastante importante.

No que diz respeito às diferenças entre quinto e nono anos, um ponto importante é que o número de escolas estaduais do primeiro ciclo do ensino fundamental é consideravelmente menor do que as de segundo ciclo. O processo de descentralização do ensino ocorrido desde meados de 1990 levou a migração da administração estadual para municipal de algumas escolas do primeiro ciclo do ensino fundamental, o que explica essa discrepância. Isso pode ter alguma relação com o fato do programa funcionar melhor para este grupo de estudantes.

Também devemos levar em consideração que a qualidade do ensino está muito aquém do desejado, e muitos alunos passam de ano sem ter alcançado um nível desejado de aprendizado para sua série. Isso acumula um déficit de aprendizado ao passar dos anos e é provável que o déficit em relação ao desejável ao chegar ao nono ano seja maior do que no quinto ano. Embora o esforço do professor possa ser maior, é mais difícil conseguir uma melhora desses alunos, principalmente quando os assuntos abordados vão ficando mais complexos.

Por fim, os alunos de quinto ano têm contato com um número bem menor de professores, o que pode contribuir para desincentivar o problema de carona dos programas de bonificação coletivos — caso do sistema estadual paulista —, já que a identificação do professor de como seu esforço pessoal vai impactar no recebimento de bônus é muito mais direta nesse caso do que no segundo ciclo do ensino fundamental. Esse último aspecto sugere, portanto, que o tratamento (exposição à política de bonificação) pode ser, na prática heterogêneo, dependendo da série considerada.

**Tabela 10.** Efeito da política de bonificação sobre indicadores escolares – 9º ano.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
<i>Contínuas</i>				
Índice socioeconômico (+)	0,10	-0,12*	0,09	-0,13*
Índice de clima escolar (-)	-0,33***	-0,05	-0,33***	-0,04
Expectativa do professor sobre futuro dos alunos (+)	0,03**	-0,02*	0,02	-0,03**
Índice de liderança do diretor (+)	0,30***	0,22**	0,37***	0,25***
Índice de uso do tempo de aula (+)	0,05	0,52***	0,06**	0,54***
Índice de cumprimento curricular	-0,01	-0,02*	-0,02*	-0,01
Formação acadêmica do professor (+)	-0,16***	-0,15***	-0,16***	-0,15***
Experiência do professor (+)	0,06***	0,16***	0,05***	0,13***
Número total de matrículas	-27,13	-41,06**	-26,84	-40,14**
Índice de Participação dos pais na escola (+)	0,18***	-0,05***	0,17***	-0,08***
Dedicação do professor (+)	0,06***	0,12***	0,05***	0,09***
<i>Discretas</i>				
Formação do diretor (<i>dummy</i> pós-graduação em gestão)	-0,06***	-0,04**	-0,02	-0,02
Experiência do diretor (<i>dummy</i> mais de 5 anos)	0,00	-0,21***	0,00	-0,21***
Alocação dos alunos em turma homogênea (por idade e rendimento)	0,01	0,13***	0,01	0,15***

Notas: Os valores correspondem aos coeficientes estimados das variáveis I_{i2009} e I_{i2011} da equação (6), onde cada indicador contido na primeira coluna da tabela foi utilizado como variável dependente dessa equação. Para detalhamento e definição dos indicadores, ver [Tavares \(2014\)](#). O sinal (+) indica que aumentos no valor do índice levam a melhoria no indicador; (-) tem a interpretação inversa.

Fonte: Elaboração própria com base nos dados da Prova Brasil de 2007, 2009 e 2011.

Cabe destacar também que políticas de bônus podem ter efeitos colaterais, como por exemplo o *teaching for the test*, prática na qual os professores focam os conteúdos exclusivamente naqueles que são cobrados nas avaliações externas, podendo inclusive a escola diminuir a carga didática de outras disciplinas que não entram no cômputo do bônus para tentar melhorar o desempenho nas disciplinas avaliadas. Assim, investigar a prática ou não desse tipo de comportamento no sistema paulista também é uma a item importante na agenda de investigação do tema.

Portanto, este trabalho apresentou alguns resultados sobre o efeito do programa de bonificação de professores sobre a proficiência, mas ainda há um espaço muito grande para discussões sobre o mesmo tema. Estudos posteriores podem ajudar a elucidar a diferença dos resultados para as diferentes séries e o decaimento do efeito no tempo, bem como compreender os canais que podem potencialmente estar mediando o impacto da política sobre o aprendizado dos alunos. Com isso, pode ser possível, então, contribuir para um melhor desenho dessa política educacional.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abadie, A., Drukker, D., Herr, J. L. & Imbens, G. W. (2004). Implementing matching estimators for average treatment effects in Stata. *The Stata Journal*, 4(3), 290–311(22). Disponível em: <http://www.stata-journal.com/article.html?article=st0072>
- Ferraz, C. (2009). Sistemas educacionais baseados em desempenho, metas de qualidade e a remuneração de professores: Os casos de Pernambuco e São paulo. In F. Veloso, S. Pessôa, R. Henriques & F. Giambiagi (Eds.), *Educação básica no Brasil: Construindo o país do futuro* (pp. 239–260). Rio de Janeiro: Elsevier.

- Ferraz, C., & Bruns, B. (2012). *Paying teachers to perform: The impact of bonus pay in Pernambuco, Brazil*. Society for Research on Educational Effectiveness.
- Figlio, D. N., & Kenny, L. W. (2007). Individual teacher incentives and student performance. *Journal of Public Economics*, 91(5-6), 901–914. doi: [10.1016/j.jpube.2006.10.001](https://doi.org/10.1016/j.jpube.2006.10.001)
- Fryer, R. G. (2013). Teacher incentives and student achievement: Evidence from New York City public schools. *Journal of Labor Economics*, 31(2), 373–407. Disponível em: <http://www.jstor.org/stable/10.1086/667757>
- Galiani, S., Gertler, P. & Scharfrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *Journal of Political Economy*, 113(1), 83–120. Disponível em: <http://www.jstor.org/stable/10.1086/426041>
- Glazerman, S., Carey, N. & McKie, A. (2009, Apr 27). *An evaluation of the Teacher Advancement Program (TAP) in Chicago: Year One Impact Report (Final Report)*. Washington, DC: Mathematica Policy Research Inc. Disponível em: <http://foundationcenter.org/educationexcellence/report.jhtml?id=fdc82500013>
- Glewwe, P., Ilias, N. & Kremer, M. (2003, May). *Teacher incentives* (NBER Working Paper N° 9671). National Bureau of Economic Research. doi: [10.3386/w9671](https://doi.org/10.3386/w9671)
- Guerra, S., & Foguel, M. (2012). *An evaluation of the Teacher Performance-Pay reforms in São Paulo and Pernambuco: A synthetic control method approach*. Encontro da SBE, 2012.
- Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal*, 113(485), F64–F98. Disponível em: <http://www.jstor.org/stable/3590139>
- Hirano, K., Imbens, G. W. & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4), 1161–1189. Disponível em: <http://www.jstor.org/stable/1555493>
- Holmstrom, B., & Milgrom, P. (1991). Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7(special issue), 24–52. doi: [10.1093/jleo/7.special_issue.24](https://doi.org/10.1093/jleo/7.special_issue.24)
- Lauen, D. L. (2011). *The effects of school wide bonuses on student achievement: Regression discontinuity evidence from North Carolina*. Society for Research on Educational Effectiveness.
- Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, 99(5), 1979–2011. doi: [10.1257/aer.99.5.1979](https://doi.org/10.1257/aer.99.5.1979)
- Muralidharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of Political Economy*, 119(1), 39–77. Disponível em: <http://www.jstor.org/stable/10.1086/659655>
- Springer, M. G., Ballou, D., Hamilton, L., Le, V., Lockwood, J. R., McCaffrey, D., ... Stecher, B. (2010). *Teacher Pay for Performance: Experimental evidence from the Project on Incentives in Teaching (POINT)*. Society for Research on Educational Effectiveness.
- São Paulo/SEE. (2009a). *Matrizes de referência para a avaliação Saesp: documento básico*. Governo do Estado de São Paulo/Secretaria de Estado da Educação. Disponível em: http://saesp.fde.sp.gov.br/2009/pdf/Saesp2008_MatrizRefAvaliacao_DocBasico_Completo.pdf
- São Paulo/SEE. (2009b, Março). *Programa de Qualidade da Escola (Nota Técnica)*. Governo do Estado de São Paulo/Secretaria de Estado da Educação. Disponível em: <http://idesp.edunet.sp.gov.br/Arquivos/NotaTecnicaPQE2008.pdf>
- Tavares, P. A. (2014). *Três ensaios em economia da educação* (Tese de Doutorado). Fundação Getulio Vargas–Escola de Economia de São Paulo.
- Vigdor, J. L. (2008). *Teacher Salary Bonuses in North Carolina* (Working Paper N° 15). National Center for Analysis of Longitudinal Data in Education Research.



A. APÊNDICE

Tabela A-1. Estimação do escore de propensão – 2007.

Variável dependente: Escola estadual paulista	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
<i>Variáveis de Alunos (Proporções)</i>				
Mãe completou a 8ª série	1,09*** (0,31)	1,01*** (0,31)	2,47*** (0,22)	2,56*** (0,22)
Possuem de 1 a 20 livros em casa	-2,45*** (0,21)	-2,43*** (0,21)	-1,84*** (0,23)	-1,84*** (0,23)
Possuem de 21 a 100 livros em casa	-2,76*** (0,28)	-2,47*** (0,27)	-3,43*** (0,25)	-3,16*** (0,24)
Possuem mais de 100 livros em casa	-2,34*** (0,34)	-2,18*** (0,34)	-2,83*** (0,4)	-2,94*** (0,4)
Fizeram o Maternal	-0,07 (0,1)	0,07 (0,1)	-0,21*** (0,07)	-0,18** (0,07)
Trabalham fora de casa	1,55*** (0,28)	1,05*** (0,27)	-0,90*** (0,16)	-0,83*** (0,16)
Mora com o pai e com a mãe	7,36*** (0,4)	7,77*** (0,4)	10,21*** (0,34)	10,60*** (0,34)
Mora com o pai ou com a mãe	5,90*** (0,45)	6,26*** (0,45)	11,00*** (0,37)	11,11*** (0,37)
<i>Variáveis de Professores (Proporções)</i>				
Salário até R\$ 1.300,00	-1,19*** (0,07)	-1,25*** (0,07)	-1,08*** (0,05)	-0,91*** (0,05)
Leciona na escola entre 3 e 5 anos	-0,79*** (0,06)	-0,78*** (0,06)	0,12*** (0,04)	0,20*** (0,03)
Leciona na escola entre 1 e 2 anos	0,40*** (0,06)	0,40*** (0,05)	0,43*** (0,04)	0,43*** (0,04)
Leciona na escola entre 6 e 9 anos	-0,64*** (0,07)	-0,64*** (0,07)	0,22*** (0,04)	0,04 (0,04)
<i>Variáveis de Diretores</i>				
Salário maior que R\$ 1.500,00	0,98*** (0,06)	1,02*** (0,05)	1,11*** (0,05)	1,16*** (0,05)
É diretor desta escola entre 2 e 4 anos	0,13** (0,05)	0,10* (0,05)	0,01 (0,04)	0,02 (0,04)
É diretor desta escola há no máximo 2 anos	0,14** (0,05)	0,12** (0,05)	0,05 (0,04)	0,05 (0,04)
É diretor desta escola entre 5 e 10 anos	0,08 (0,06)	0,06 (0,05)	-0,01 (0,04)	-0,01 (0,04)
<i>Variáveis de Escola</i>				
Ln do tamanho da Escola	0,35*** (0,02)	0,34*** (0,02)	0,42*** (0,02)	0,40*** (0,02)
<i>Variáveis de Tendência</i>				
Média de proficiência da Escola em 2007	0,01*** (0,00)	0,00** (0,00)	0,00* (0,00)	0,00*** (0,00)
Média da Variação da proficiência da escola entre 2005 e 2007	-4,82*** (0,24)	-2,41*** (0,23)	-1,45*** (0,24)	0,12 (0,26)
Constante	-10,90*** (0,48)	-9,81*** (0,47)	-12,21*** (0,43)	-10,90*** (0,42)

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A-2. Diferença de médias das variáveis explicativas – Quinto ano – Português.

	Antes do Matching			Depois do Matching		
	Tratamento	Controle	$p > t$	Tratamento	Controle	$p > t$
<i>Variáveis de Alunos (Proporções)</i>						
Mãe completou a 8ª série	0,140	0,125	0,00	0,140	0,142	0,29
Possuem de 1 a 20 livros em casa	0,524	0,517	0,03	0,525	0,524	0,69
Possuem de 21 a 100 livros em casa	0,159	0,158	0,82	0,159	0,157	0,39
Possuem mais de 100 livros em casa	0,092	0,093	0,55	0,092	0,092	0,82
Fizeram o Maternal	0,343	0,336	0,12	0,343	0,340	0,51
Trabalham fora de casa	0,127	0,131	0,03	0,126	0,127	0,58
Mora com o pai e com a mãe	0,618	0,525	0,00	0,618	0,625	0,04
Mora com o pai ou com a mãe	0,325	0,339	0,00	0,325	0,321	0,20
<i>Variáveis de Professores (Proporções)</i>						
Salário até R\$ 1.300,00	0,079	0,309	0,00	0,081	0,076	0,43
Leciona na escola entre 3 e 5 anos	0,119	0,225	0,00	0,122	0,125	0,68
Leciona na escola entre 1 e 2 anos	0,202	0,129	0,00	0,197	0,198	0,92
Leciona na escola entre 6 e 9 anos	0,108	0,184	0,00	0,111	0,099	0,09
<i>Variáveis de Diretores (Proporções)</i>						
Salário maior que R\$ 1.500,00	0,959	0,593	0,00	0,958	0,937	0,01
É diretor desta escola entre 2 e 4 anos	0,328	0,353	0,06	0,329	0,328	0,94
É diretor desta escola há no máximo 2 anos	0,331	0,322	0,48	0,324	0,333	0,60
É diretor desta escola entre 5 e 10 anos	0,250	0,232	0,12	0,254	0,250	0,83
<i>Variáveis de Escola</i>						
Ln do tamanho da Escola	6,58	6,32	0,00	6,57	6,55	0,40
<i>Variáveis de Tendência</i>						
Média da Escola na proficiência em 2007	177,488	172,496	0,00	177,70782	177,8395	0,83
Média da Variação da proficiência da escola entre 2005 e 2007	-0,03401	0,00513	0,00	-0,03101	-0,02526	0,03

**Tabela A-3.** Diferença de médias das variáveis explicativas – Quinto ano – Matemática.

	Antes do Matching			Depois do Matching		
	Tratamento	Controle	$p > t$	Tratamento	Controle	$p > t$
<i>Variáveis de Alunos (Proporções)</i>						
Mãe completou a 8ª série	0,140	0,125	0,00	0,140	0,140	0,98
Possuem de 1 a 20 livros em casa	0,524	0,517	0,03	0,525	0,524	0,80
Possuem de 21 a 100 livros em casa	0,159	0,158	0,81	0,159	0,159	0,90
Possuem mais de 100 livros em casa	0,092	0,093	0,51	0,092	0,092	0,76
Fizeram o Maternal	0,342	0,336	0,13	0,343	0,340	0,69
Trabalham fora de casa	0,127	0,131	0,03	0,127	0,127	0,90
Mora com o pai e com a mãe	0,618	0,525	0,00	0,618	0,621	0,32
Mora com o pai ou com a mãe	0,325	0,339	0,00	0,325	0,323	0,42
<i>Variáveis de Professores (Proporções)</i>						
Salário até R\$ 1.300,00	0,076	0,308	0,00	0,076	0,075	0,75
Leciona na escola entre 3 e 5 anos	0,118	0,224	0,00	0,118	0,122	0,64
Leciona na escola entre 1 e 2 anos	0,201	0,128	0,00	0,200	0,198	0,85
Leciona na escola entre 6 e 9 anos	0,106	0,183	0,00	0,107	0,099	0,29
<i>Variáveis de Diretores (Proporções)</i>						
Salário maior que R\$ 1.500,00	0,959	0,593	0,00	0,959	0,958	0,85
É diretor desta escola entre 2 e 4 anos	0,328	0,353	0,06	0,329	0,327	0,94
É diretor desta escola há no máximo 2 anos	0,331	0,322	0,46	0,330	0,328	0,94
É diretor desta escola entre 5 e 10 anos	0,249	0,232	0,13	0,250	0,247	0,86
<i>Variáveis de Escola</i>						
Ln do tamanho da Escola	6,5835	6,3242	0,00	6,58214	6,57251	0,66
<i>Variáveis de Tendência</i>						
Média da Escola na proficiência em 2007	194,51	194,01	0,00	194,55	194,90	0,58
Média da Variação da proficiência da escola entre 2005 e 2007	0,040	0,060	0,00	0,040	0,044	0,16

Tabela A-4. Diferença de médias das variáveis explicativas – Nono ano – Português.

	Antes do Matching			Depois do Matching		
	Tratamento	Controle	$p > t$	Tratamento	Controle	$p > t$
<i>Variáveis de Alunos (Proporções)</i>						
Mãe completou a 8ª série	0,184	0,154	0,00	0,183	0,185	0,29
Possuem de 1 a 20 livros em casa	0,602	0,589	0,00	0,601	0,603	0,42
Possuem de 21 a 100 livros em casa	0,184	0,195	0,00	0,185	0,183	0,46
Possuem mais de 100 livros em casa	0,056	0,065	0,00	0,056	0,056	0,59
Fizeram o Maternal	0,419	0,425	0,17	0,421	0,418	0,54
Trabalham fora de casa	0,182	0,209	0,00	0,182	0,181	0,69
Mora com o pai e com a mãe	0,618	0,596	0,00	0,617	0,620	0,35
Mora com o pai ou com a mãe	0,321	0,294	0,00	0,321	0,319	0,56
<i>Variáveis de Professores (Proporções)</i>						
Salário até R\$ 1.300,00	0,065	0,268	0,00	0,069	0,061	0,09
Leciona na escola entre 3 e 5 anos	0,207	0,209	0,73	0,208	0,207	0,97
Leciona na escola entre 1 e 2 anos	0,181	0,133	0,00	0,175	0,168	0,37
Leciona na escola entre 6 e 9 anos	0,191	0,181	0,13	0,190	0,194	0,63
<i>Variáveis de Diretores (Proporções)</i>						
Salário maior que R\$ 1.500,00	0,973	0,692	0,00	0,971	0,967	0,39
É diretor desta escola entre 2 e 4 anos	0,307	0,318	0,27	0,308	0,301	0,58
É diretor desta escola há no máximo 2 anos	0,344	0,355	0,25	0,343	0,341	0,84
É diretor desta escola entre 5 e 10 anos	0,246	0,236	0,24	0,245	0,254	0,43
<i>Variáveis de Escola</i>						
Ln do tamanho da Escola	6,913	6,701	0,00	6,897	6,866	0,02
<i>Variáveis de Tendência</i>						
Média da Escola na proficiência em 2007	231,27	228,05	0,00	231,47	232,22	0,06
Média da Variação da proficiência da escola entre 2005 e 2007	0,011	0,020	0,00	0,012	0,013	0,69

**Tabela A-5.** Diferença de médias das variáveis explicativas – Nono ano – Matemática.

	Antes do Matching			Depois do Matching		
	Tratamento	Controle	$p > t$	Tratamento	Controle	$p > t$
<i>Variáveis de Alunos (Proporções)</i>						
Mãe completou a 8ª série	0,184	0,154	0,00	0,183	0,186	0,10
Possuem de 1 a 20 livros em casa	0,602	0,590	0,00	0,602	0,603	0,75
Possuem de 21 a 100 livros em casa	0,183	0,195	0,00	0,184	0,183	0,36
Possuem mais de 100 livros em casa	0,056	0,065	0,00	0,056	0,056	0,76
Fizeram o Maternal	0,419	0,424	0,24	0,421	0,427	0,23
Trabalham fora de casa	0,182	0,209	0,00	0,183	0,180	0,17
Mora com o pai e com a mãe	0,618	0,597	0,00	0,619	0,619	0,88
Mora com o pai ou com a mãe	0,321	0,294	0,00	0,320	0,320	0,78
<i>Variáveis de Professores (Proporções)</i>						
Salário até R\$ 1.300,00	0,072	0,268	0,00	0,076	0,071	0,36
Leciona na escola entre 3 e 5 anos	0,246	0,227	0,01	0,245	0,243	0,86
Leciona na escola entre 1 e 2 anos	0,180	0,126	0,00	0,171	0,160	0,2
Leciona na escola entre 6 e 9 anos	0,156	0,175	0,01	0,159	0,162	0,73
<i>Variáveis de Diretores (Proporções)</i>						
Salário maior que R\$ 1.500,00	0,974	0,692	0,00	0,972	0,969	0,58
É diretor desta escola entre 2 e 4 anos	0,307	0,317	0,30	0,306	0,292	0,26
É diretor desta escola há no máximo 2 anos	0,343	0,356	0,19	0,341	0,343	0,89
É diretor desta escola entre 5 e 10 anos	0,247	0,235	0,18	0,248	0,249	0,93
<i>Variáveis de Escola</i>						
Ln do tamanho da Escola	6,9134	6,7002	0,00	6,898	6,8712	0,05
<i>Variáveis de Tendência</i>						
Média da Escola na proficiência em 2007	242,2	240,04	0,00	242,64	242,98	0,42
Média da Variação da proficiência da escola entre 2005 e 2007	0,001	0,004	0,042	0,002	0,000	0,42

Tabela A-6. Resultado do Diferenças em Diferenças com reponderação.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
EE ^{SP}	-2,36*** (0,60)	-2,59*** (0,57)	-1,83*** (0,41)	-2,84*** (0,41)
2009	10,97*** (0,40)	13,94*** (0,41)	11,46*** (0,41)	2,61*** (0,43)
2011	17,89*** (0,41)	19,79*** (0,43)	12,02*** (0,41)	6,89*** (0,43)
Interação 2009	2,36*** (0,83)	6,01*** (0,87)	-2,44*** (0,61)	-2,09*** (0,62)
Interação 2011	-1,41 (0,89)	1,54* (0,93)	-1,67*** (0,61)	-3,74*** (0,63)
Constante	177,99*** (0,27)	195,02*** (0,26)	231,21*** (0,27)	242,18*** (0,29)
Observações	69.636	69.615	52.689	52.686
R ²	0,1553	0,1584	0,1066	0,039

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A-7. Resultado do Diferenças em Diferenças por Município.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
EE ^{SP}	-3,30*** (0,48)	-4,41*** (0,50)	-2,57*** (0,64)	-3,91*** (0,64)
2009	9,03*** (0,46)	12,46*** (0,50)	9,47*** (0,77)	0,15 (0,76)
2011	14,09*** (0,48)	15,16*** (0,53)	12,85*** (0,75)	5,39*** (0,78)
Interação 2009	4,02*** (0,67)	7,25*** (0,73)	-1,14 (0,94)	-0,19 (0,92)
Interação 2011	2,34*** (0,69)	6,16*** (0,75)	-2,51*** (0,91)	-2,59*** (0,94)
Constante	184,27*** (2,93)	203,63*** (3,43)	256,15*** (3,34)	264,75*** (3,72)
Dummy Municípios	Sim	Sim	Sim	Sim
Observações	9.530	9.527	5.869	5.869
R ²	0,37	0,428	0,26	0,259

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

**Tabela A-8.** Diferenças em Diferenças: Estaduais paulistas × Municipais paulistas.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
EE ^{SP}	0,83 (0,70)	0,23 (0,72)	0,51 (0,81)	0,32 (0,87)
2009	10,60*** (0,74)	15,53*** (0,82)	8,56*** (0,90)	-0,47 (0,96)
2011	16,04*** (0,74)	18,45*** (0,85)	12,75*** (0,88)	5,41*** (0,98)
Interação 2009	1,26 (1,00)	4,00*** (1,09)	-0,35 (1,18)	0,32 (1,26)
Interação 2011	-1,49 (1,00)	1,72 (1,12)	-3,86*** (1,15)	-3,36*** (1,28)
Constante	179,41*** (0,51)	197,18*** (0,53)	233,63*** (0,61)	245,84*** (0,66)
Observações	5.640	5.814	3.960	3.900
R ²	0,1489	0,1983	0,089	0,0154

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A-9. Diferenças em Diferenças: Estaduais paulistas × Resto do Brasil (Estaduais).

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
EE ^{SP}	0,69 (0,65)	0,37 (0,67)	-0,46 (0,43)	-0,08 (0,46)
2009	9,90*** (0,73)	12,07*** (0,81)	11,02*** (0,49)	2,78*** (0,52)
2011	17,44*** (0,76)	18,40*** (0,81)	10,53*** (0,49)	5,61*** (0,54)
Interação 2009	2,59*** (0,96)	7,71*** (1,04)	-2,21*** (0,64)	-2,11*** (0,67)
Interação 2011	-2,42** (0,98)	2,16** (1,04)	-0,3 (0,64)	-2,40*** (0,69)
Constante	176,99*** (0,49)	194,19*** (0,52)	232,32*** (0,33)	243,23*** (0,36)
Observações	6.630	7.122	13.338	13.554
R ²	0,1491	0,1853	0,0906	0,0154

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A-10. Resultado do Diferenças em Diferenças: Estaduais paulistas × Escolas estaduais do restante do Sudeste.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
EE ^{SP}	0,01 (0,62)	-0,29 (0,64)	-0,45 (0,46)	0,43 (0,50)
2009	12,52*** (0,68)	16,86*** (0,75)	11,54*** (0,52)	3,77*** (0,57)
2011	17,47*** (0,69)	20,60*** (0,75)	13,98*** (0,53)	9,01*** (0,59)
Intereção 2009	-0,16 (0,9)	2,88*** (0,99)	-2,61*** (0,69)	-3,14*** (0,74)
Interação 2011	-2,69*** (0,91)	-0,19 (0,99)	-3,64*** (0,69)	-6,09*** (0,76)
Constante	178,68*** (0,46)	195,94*** (0,48)	233,15*** (0,35)	243,84*** (0,39)
Observações	7.386	7.596	11.490	11.040
R ²	0,1579	0,2052	0,1152	0,0335

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A-11. Resultado do Diferenças em Diferenças: Estaduais paulistas × Estaduais Rio de Janeiro e Espírito Santo.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
EE ^{SP}	-0,46 (0,63)	-0,22 (0,65)	0,16 (0,58)	-0,58 (0,64)
2009	10,67*** (0,68)	14,85*** (0,76)	10,20*** (0,64)	1,97*** (0,70)
2011	16,00*** (0,69)	18,95*** (0,77)	11,63*** (0,66)	7,01*** (0,74)
Interação 2009	1,58* (0,91)	4,85*** (0,99)	-1,42 (0,86)	-1,13 (0,92)
Interação 2011	-1,42 (0,92)	1,46 (1,01)	-1,26 (0,87)	-3,38*** (0,96)
Constante	179,53*** (0,47)	196,15*** (0,50)	232,30*** (0,43)	243,72*** (0,49)
Observações	7.014	7.332	7.242	7.170
R ²	0,1462	0,1958	0,0922	0,0239

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.



Tabela A-12. Resultado do Diferenças em Diferenças: Estaduais paulistas × Escolas do Brasil, excluindo Redes que também adotaram bônus (RJ, PE, GO, ES e MG).

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
EE ^{SP}	-1,06*	-0,71	0,27	0,32
	(0,60)	(0,63)	(0,39)	(0,41)
2009	10,85***	13,95***	11,65***	2,93***
	(0,67)	(0,75)	(0,44)	(0,47)
2011	17,76***	20,31***	12,40***	6,72***
	(0,68)	(0,77)	(0,44)	(0,49)
Interação 2009	1,85**	5,92***	-2,80***	-2,28***
	(0,87)	(0,97)	(0,57)	(0,61)
Interação 2011	-2,54***	0,41	-2,26***	-3,59***
	(0,88)	(0,98)	(0,57)	(0,62)
Constante	178,77***	195,34***	231,25***	242,24***
	(0,45)	(0,48)	(0,30)	(0,32)
Observações	8.316	8.472	16.356	16.500
R ²	0,1533	0,1884	0,1032	0,0196

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A-13. Diferenças triplas: Estadual SP vs Municipal SP × Estadual Br vs Municipal Br.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
Dummy de UF (SP)	16,16*** (0,30)	18,30*** (0,35)	9,75*** (0,50)	11,07*** (0,55)
Escola Estadual (EE)	5,13*** (0,21)	4,95*** (0,23)	3,42*** (0,24)	4,24*** (0,26)
SP·EE	-12,68*** (0,48)	-14,76*** (0,52)	-6,13*** (0,56)	-8,59*** (0,61)
2009	9,79*** (0,18)	11,68*** (0,21)	10,92*** (0,29)	2,62*** (0,31)
2011	16,81*** (0,19)	18,07*** (0,22)	12,62*** (0,29)	8,36*** (0,33)
SP·2009	-1,28*** (0,43)	1,56*** (0,51)	-1,99*** (0,73)	-2,76*** (0,79)
SP·2011	-3,58*** (0,44)	-2,32*** (0,52)	-0,34 (0,71)	-3,03*** (0,80)
EE·2009	0,02 (0,33)	0,83** (0,38)	0,34 (0,36)	0,22 (0,39)
EE·2011	-0,07 (0,35)	-0,04 (0,40)	-1,11*** (0,37)	-1,34*** (0,41)
EE·SP	-12,68*** (0,48)	-14,76*** (0,52)	-6,13*** (0,56)	-8,59*** (0,61)
EE·SP·2009	4,06*** (0,70)	5,43*** (0,80)	-0,33 (0,83)	0,63 (0,89)
EE·SP·2011	2,37*** (0,72)	4,86*** (0,82)	-0,78 (0,82)	-0,61 (0,90)
Constante	167,87*** (0,12)	185,14*** (0,13)	224,10*** (0,19)	235,31*** (0,21)
R^2	0,1553	0,1575	0,0973	0,0347
Observações	104.015	104.015	78.181	78.181

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

**Tabela A-14.** Diferenças triplas: Estadual SP vs Municipal SP × Estadual SE vs Municipal SE.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
SP	7,96*** (0,37)	9,17*** (0,41)	-0,16 (0,59)	1,19* (0,66)
EE	3,91*** (0,42)	3,97*** (0,46)	-2,23*** (0,46)	0,00 (0,54)
SP-EE	-11,46*** (0,60)	-13,79*** (0,65)	-0,47 (0,69)	-4,35*** (0,77)
2009	15,78*** (0,35)	19,04*** (0,40)	11,30*** (0,54)	3,74*** (0,60)
2011	22,41*** (0,35)	25,87*** (0,40)	16,26*** (0,53)	12,98*** (0,61)
SP-2009	-7,27*** (0,53)	-5,79*** (0,61)	-2,37*** (0,86)	-3,88*** (0,95)
SP-2011	-9,18*** (0,53)	-10,12*** (0,61)	-3,97*** (0,84)	-7,65*** (0,95)
EE-2009	-1,69*** (0,64)	0,54 (0,74)	2,32*** (0,68)	2,23*** (0,77)
EE-2011	-3,44*** (0,65)	-4,01*** (0,72)	-1,87*** (0,69)	-2,41*** (0,80)
EE-SP	-11,46*** (0,60)	-13,79*** (0,65)	-0,47 (0,69)	-4,35*** (0,77)
EE-SP-2009	5,78*** (0,89)	5,71*** (1,02)	-2,32** (1,01)	-1,37 (1,11)
EE-SP-2011	5,74*** (0,90)	8,83*** (1,02)	-0,01 (1,00)	0,45 (1,13)
Constante	176,07*** (0,24)	194,27*** (0,26)	234,03*** (0,37)	245,20*** (0,42)
R^2	0,1786	0,1946	0,1269	0,0697
Observações	37.566	37.566	30.244	30.244

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A-15. Diferenças triplas: Estadual SP vs Municipal SP × Estadual RJ + ES vs Municipal RJ + ES.

Variáveis	Quinto ano		Nono ano	
	Português	Matemática	Português	Matemática
SP	6,16*** (0,40)	9,12*** (0,46)	0,91 (0,65)	4,60*** (0,72)
EE	-4,03*** (0,67)	-3,96*** (0,71)	-8,81*** (0,65)	-8,32*** (0,72)
SP-EE	-3,51*** (0,79)	-5,84*** (0,85)	6,11*** (0,83)	3,97*** (0,91)
2009	10,58*** (0,45)	12,82*** (0,49)	11,26*** (0,67)	3,73*** (0,72)
2011	16,83*** (0,45)	21,34*** (0,51)	15,07*** (0,67)	12,86*** (0,76)
SP·2009	-2,07*** (0,59)	0,42 (0,68)	-2,32** (0,95)	-3,87*** (1,03)
SP·2011	-3,60*** (0,60)	-5,59*** (0,70)	-2,79*** (0,94)	-7,54*** (1,05)
EE·2009	-4,03*** (1,04)	-3,19*** (1,14)	1,19 (1,00)	2,73*** (1,06)
EE·2011	-6,29*** (1,12)	-6,00*** (1,21)	-6,28*** (1,03)	-5,06*** (1,14)
EE·SP	-3,51*** (0,79)	-5,84*** (0,85)	6,11*** (0,83)	3,97*** (0,91)
EE·SP·2009	8,12*** (1,21)	9,45*** (1,34)	-1,18 (1,25)	-1,87 (1,33)
EE·SP·2011	8,58*** (1,28)	10,82*** (1,41)	4,38*** (1,26)	3,10** (1,40)
Constante	177,86*** (0,29)	194,32*** (0,33)	232,95*** (0,46)	241,78*** (0,52)
R ²	0,1585	0,194	0,1342	0,0649
Observações	25.845	25.845	20.766	20.766

Notas: Erro Padrão robusto a heterocedasticidade entre parênteses.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.



Figura A-1. Distribuição de escolas de acordo com a fração de cumprimento das metas da política de bônus da rede estadual de São Paulo entre 2008 e 2012.

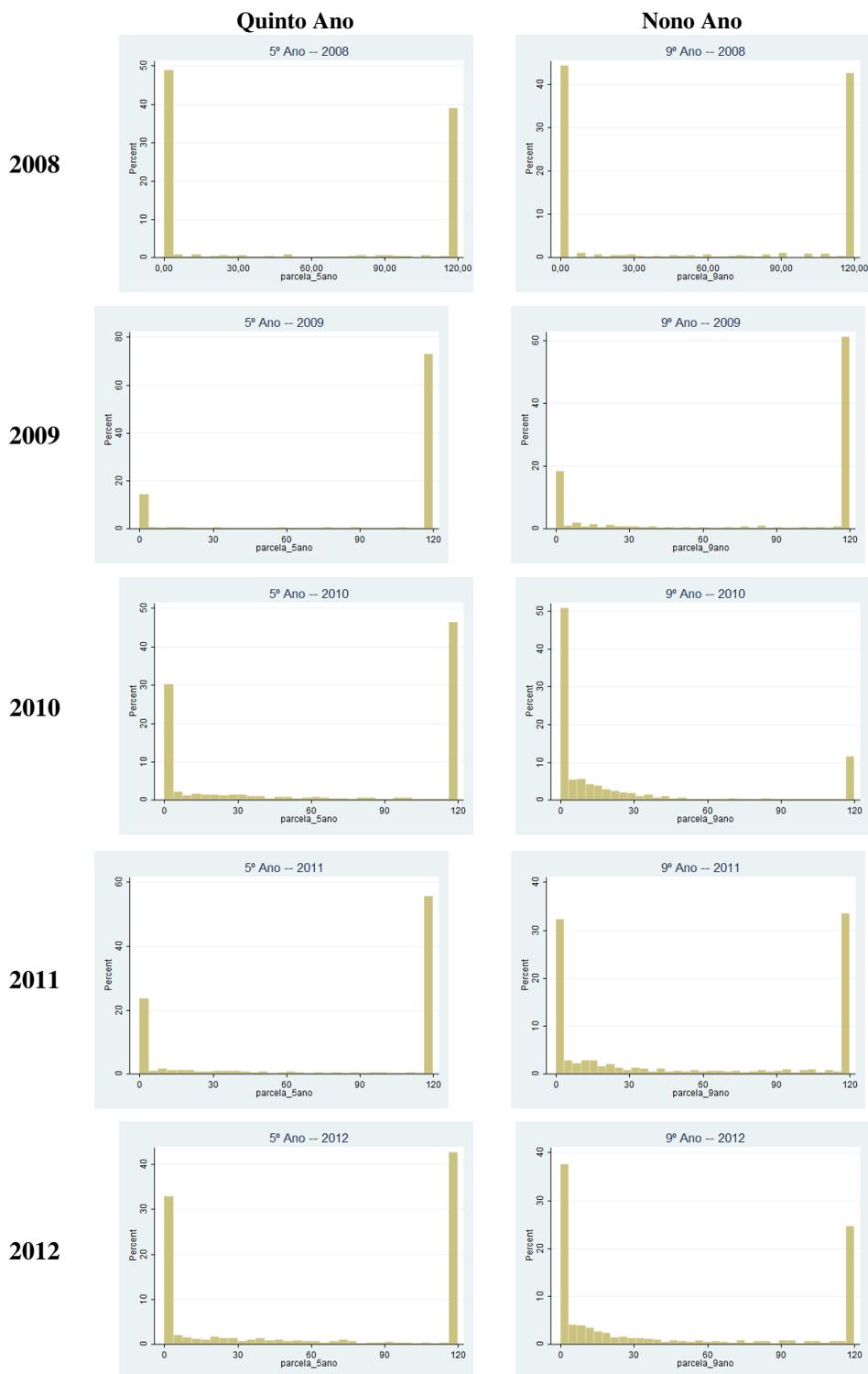


Figura A-2. Evolução das notas entre 2005 e 2011 para escolas estaduais e escolas municipais do estado de São Paulo.

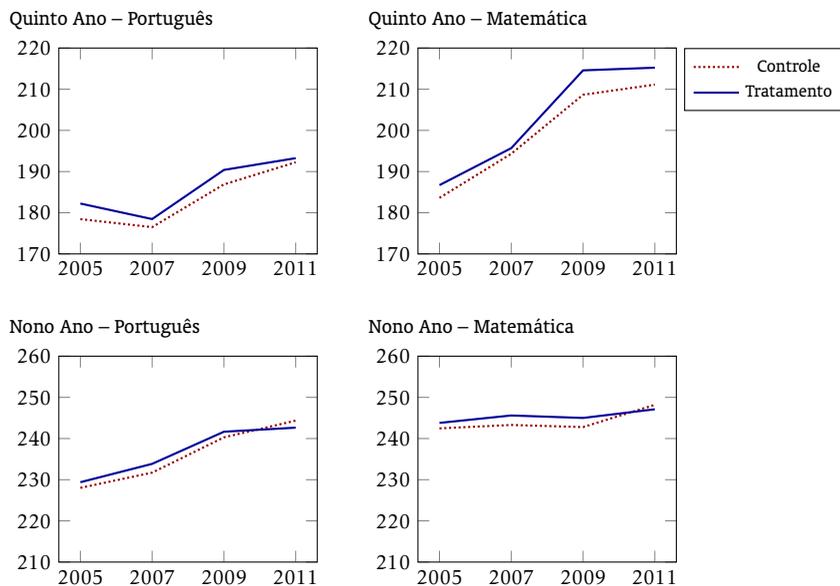


Figura A-3. Evolução das notas entre 2005 e 2011 para escolas estaduais paulistas e escolas estaduais do resto do Brasil.

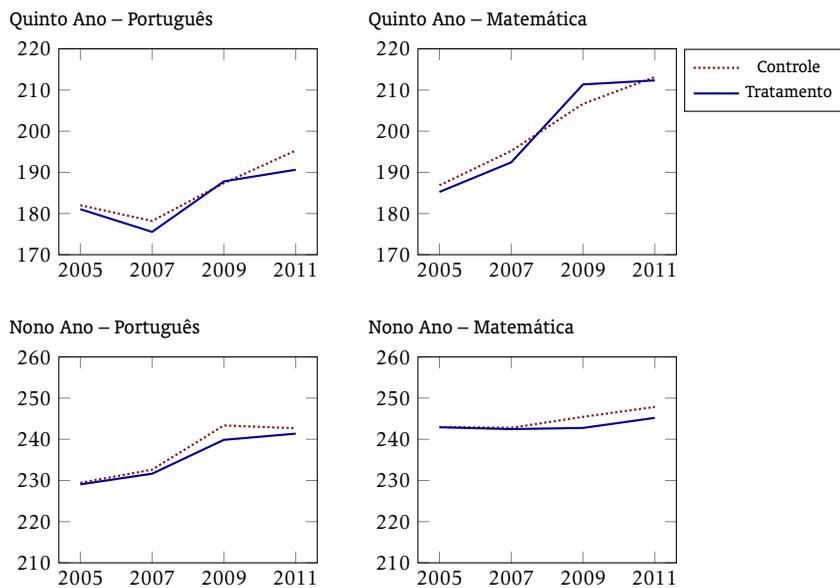




Figura A-4. Evolução das notas entre 2005 e 2011 para escolas estaduais paulistas e escolas estaduais do Sudeste.

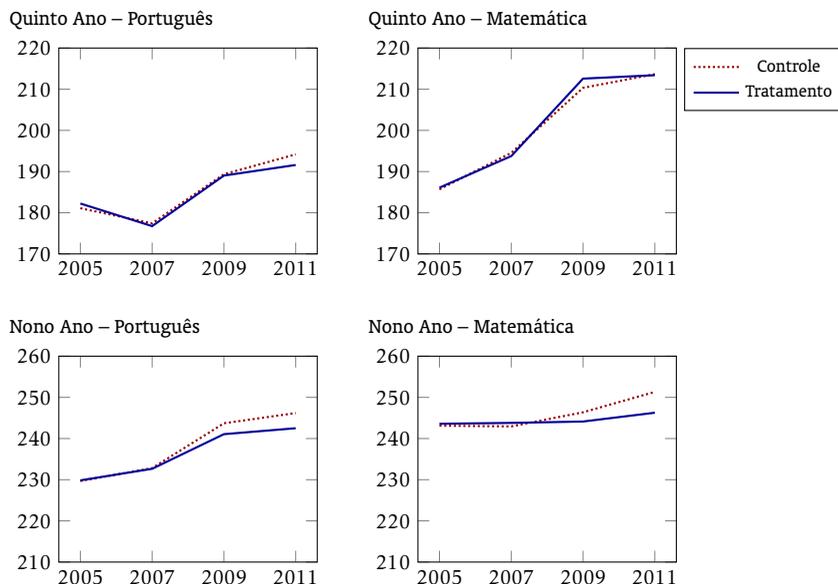


Figura A-5. Evolução das notas entre 2005 e 2011 para escolas estaduais paulistas e escolas estaduais do Rio de Janeiro e do Espírito Santo.

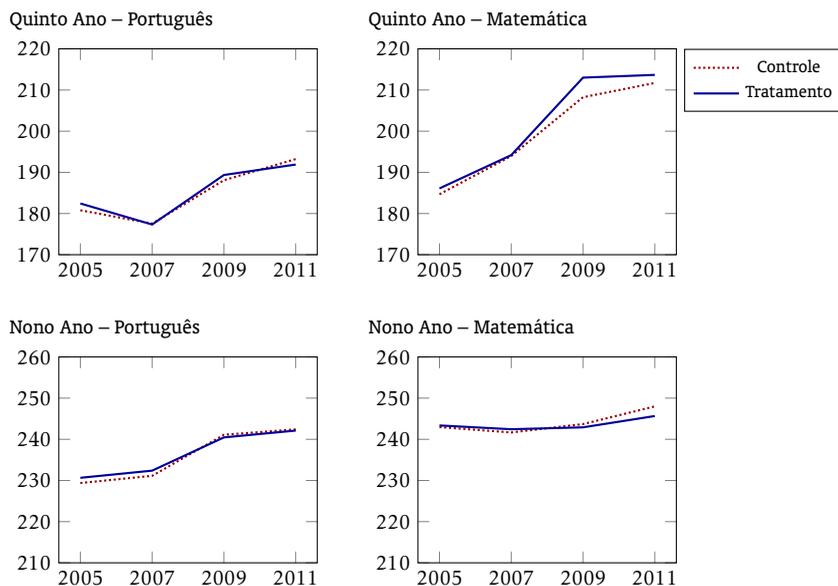


Figura A-6. Evolução das notas entre 2005 e 2011 para escolas estaduais paulistas e escolas do restante do Brasil, exceto redes estaduais de MG, RJ, ES, GO, CE e PE.

