

EFEITO DO PAGAMENTO DE BÔNUS AOS PROFESSORES SOBRE A PROFICIÊNCIA ESCOLAR NO ESTADO DE SÃO PAULO

Cláudia Hiromi Oshiro ¹
Luiz Guilherme Scorzafave ²

Resumo

Um dos maiores desafios brasileiros é a melhora do sistema educacional. Embora um avanço considerável tenha ocorrido no que diz respeito ao acesso à escola nas últimas duas décadas, a qualidade do ensino ainda deixa muito a desejar. Em busca de melhora da qualidade, diversas políticas têm sido adotadas nas diferentes instâncias de governo. Neste trabalho, será avaliado o programa implementado no estado de São Paulo a partir de 2008 que paga uma bonificação aos professores e funcionários da escola que cumpre as metas pré-estabelecidas pela Secretaria de Educação, sendo que a melhoria das notas dos alunos em testes padronizados é elemento importante do sistema. Os métodos de *Propensity Score Matching* e Diferenças em Diferenças foram utilizados para medir o impacto do programa sobre a proficiência dos alunos das escolas estaduais paulistas. Foram analisados os efeitos sobre as provas de Matemática e Língua Portuguesa para a quarta e oitava série do ensino fundamental. Resultados significativos e positivos foram encontrados apenas para as notas da quarta série. O efeito para Matemática foi de cerca de 6,4 pontos e o de Língua Portuguesa em cerca de 3,7 pontos na escala SAEB.

Abstract

One of the biggest Brazil's challenges is the improvement of educational system. Although a huge progress has occurred concerning the access to school in the last two decades, the quality of education is yet far from ideal. To change this figure, various programs have been adopted by government. In this paper, we evaluate the program implemented in São Paulo state since 2008 that pay a bonus to teachers and other workers of schools that achieve the targets previously established by the government and the improvement in standardized tests is a important element of the system. Propensity Score Matching and Differences-in-Differences were used to measure the impact of program on the average scores of schools. We have analyzed the effects on Mathematics and Portuguese for the fourth and eighth grade of elementary school. Significant and positive results were found only in fourth grade.

Palavras-Chave: *Accountability*, Bônus, Educação, Professores

Classificação JEL: I28, I25

Área ANPEC: 11 - Economia Social e Demografia Econômica

¹ Mestranda em Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo (FEA-RP/USP). Email: claudiahiromi@usp.br

² Professor do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo (FEA-RP/USP). Email: scorza@usp.br

EFEITO DO PAGAMENTO DE BÔNUS AOS PROFESSORES SOBRE A PROFICIÊNCIA ESCOLAR

1 – Introdução

A melhora do sistema educacional é um dos grandes desafios para o crescimento da economia brasileira. É inegável o avanço ocorrido na educação nas duas últimas décadas, tendo sido a quase universalização do ensino fundamental uma de suas maiores conquistas. No entanto, a formação de capital humano requer educação de qualidade e, nesse sentido, o sistema brasileiro ainda deixa a desejar. Embora os dados do Sistema de Avaliação da Educação Básica (SAEB) apontem crescimento das notas nas últimas avaliações – após período de queda entre 1995 e 2001³ -, o nível educacional 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 OCDE.

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

No Brasil, a criação de métodos de avaliação do ensino a partir da década de 1990 tornou possível a elaboração de mecanismos de responsabilização. Um importante passo foi a reformulação do SAEB, em 1995. Com o passar dos anos, houve uma evolução considerável dos sistemas de avaliação nas esferas estaduais e municipais. Hoje, após algumas revisões, muitas metodologias tornaram-se compatíveis, o que permite a comparação de resultados ao longo do tempo e entre pesquisas.

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 da notas pode gerar pressão por parte dos pais e da comunidade como um todo sobre as instituições com piores resultados. Também é comum oferecer premiações ou impor punições condicionadas ao desempenho em exames (Veloso, 2009).

Diante deste contexto, o objetivo deste trabalho é investigar se o programa de bonificação de professores e funcionários de escolas da rede estadual de São Paulo implementado em 2008, ao responsabilizar os agentes escolares pelo desempenho dos alunos, teve efeito sobre as notas dos alunos nos testes padronizados.

Para este estudo utilizaremos a metodologia de *Propensity Score Matching* e de Diferenças em Diferenças. Com esta abordagem, pode-se encontrar um grupo de controle (não afetado pela política de bônus) que seja semelhante ao grupo tratamento (escolas afetadas pelo programa) com base em características observáveis e não observáveis (constantes no tempo). Com isso, pode-se inferir se as escolas sujeitas ao programa tiveram melhora de desempenho devido a esta intervenção. Ou seja, avaliaremos se o programa de bonificação gera incentivos capazes de mudar a conduta do professor e de outros agentes escolares de forma que o aprendizado na escola melhore de uma forma geral levando, portanto, a um aumento da proficiência nos testes.

Poucos estudos empíricos nesta área foram realizados no Brasil. Uma das razões é que várias políticas adotadas ainda são muito recentes, a ponto de existirem poucos dados suficientes para estudos mais aprofundados. Diante disso, este trabalho pretende preencher essa lacuna.

O restante do trabalho se divide em mais sete seções. A próxima seção faz uma revisão de importantes artigos sobre políticas de responsabilização em geral. A terceira explica o funcionamento do programa de bônus paulista. A quarta e a quinta seções apresentam a base de dados e a metodologia, respectivamente. Os resultados estão na sexta seção e, por fim, a conclusão na última seção.

³ De 1995 a 2001 houve queda das notas de Português para a quarta série e de Matemática para quarta e oitava. As notas de Português da oitava série apresentaram queda de 1995 até 1999. Em geral, os resultados apresentaram certa estabilidade até 2005 e depois melhora em 2007 e 2009.

2 – Revisão da literatura

As reformas educacionais nas últimas décadas têm como um de seus pilares a responsabilização de agentes escolares pelo aprendizado dos alunos. Uma das reformas mais importantes ocorreu nos Estados Unidos, em 2001, com o *No Child Left Behind Act* (NCLB) que passou a focar a política educacional na questão do desempenho com ênfase em políticas de incentivo.

Programas de *accountability*⁴ foram ganhando espaço na medida em que as velhas práticas de prover insumos e recursos para as escolas se mostraram insuficientes para garantir uma boa 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.

O debate sobre mecanismos de melhoria de qualidade dos docentes e seus impactos sobre a proficiência do aluno é um dos temas mais importantes da literatura. Dolton e Gutierrez (2011) discutiram as diferenças salariais de professores entre 39 países e sua consequência sobre os resultados dos estudantes. A remuneração dada a esses profissionais reflete, primeiramente, a importância que o país dá à educação. Em países com altos níveis educacionais, os professores estão no topo da distribuição salarial entre diversas categorias profissionais.

Usando dados em painel de 12 anos, foi encontrado que cerca de 10% de aumento nos salários levam a uma melhora de desempenho entre 5% e 10%. Salários mais altos funcionam como uma forma de melhorar a qualidade de ensino no longo prazo, uma vez que incentivam maior esforço dos professores para se manterem no emprego e atraem melhores profissionais, gerando um mercado de trabalho mais competitivo e, desta forma, aumentando a habilidade média dos trabalhadores da área.

No entanto, mudar a estrutura salarial de uma determinada carreira é uma tarefa de prazo mais longo. Dentro do contexto de valorização da profissão de docente, vários países têm adotado políticas de pagamento por mérito. Este assunto não chega a ser um consenso entre os estudiosos de educação. Milgrom e Homstrom (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. Os críticos dessa prática argumentam que o incentivo levaria o professor a se concentrar apenas 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 que não são medidas nas provas.

Um dos programas norte-americanos de bonificação de professores mais antigos é o da Carolina do Norte. Desde 1996, o estado remunera os professores de acordo com o desempenho da escola. Vigdor (2008) investigou o programa e observou melhora nos exames de Matemática. No entanto, o autor ainda faz ressalvas quanto à estrutura de incentivos adotada – baseada nos resultados gerais da escola –, que pode levar ao problema do *free-rider*.

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 dos professores e desempenho dos alunos nos Estados Unidos. Eles mostraram que escolas que oferecem pagamento por mérito tiveram 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 tais políticas. Neste caso, o aumento das notas seria um resultado espúrio. A metodologia adotada no trabalho, porém, não possibilitou tal investigação.

Outro estudo nessa área foi realizado por Lavy (2009) em Israel. O programa experimental recompensava individualmente os professores - com bônus em dinheiro - pela melhora do desempenho dos estudantes em exames de matrícula no ensino médio. O autor utilizou a metodologia de Regressão

⁴ O termo *accountability* tem sido traduzido como transparência e responsabilização, entre outros. Na falta de concordância sobre a melhor tradução, foi mantido o termo em inglês (Fernandes e Gremaud (2009)). Além disso, o sistema de *accountability* é um conceito mais amplo do que apenas a responsabilização. Ele exige uma série de outros parâmetros.

Descontínua e de Diferenças em Diferenças e encontrou melhora nas notas. O autor qualifica incentivos aos docentes como um “método promissor de melhora da qualidade escolar”. A ressalva com relação a esse trabalho é que o experimento, que teve duração de apenas um ano, não permitiu estudar efeitos dos incentivos em outras coortes e identificar efeitos de longo prazo.

Também em caráter temporário, foi estruturado um programa de premiação de professores no Quênia condicionado ao desempenho dos estudantes. Ao fazer a análise de tal projeto, Glewwe et al. (2003) encontraram maiores ganhos de notas em escolas participantes durante o período do programa. No entanto, não verificaram mudanças em atitudes dos professores, seja em relação à maior presença em sala de aula ou nos métodos pedagógicos. A única mudança ocorreu na alocação de tempo para treinamento para a prova, o que deve explicar o aumento das notas. Os autores atribuem à natureza temporária da política os resultados não muito animadores.

Atkinson et. al. (2009) analisaram efeitos de um programa inglês usando dados longitudinais e método de Diferenças em Diferenças. O programa remunera o professor de acordo com o seu desempenho em cinco áreas. Uma delas é a melhora das notas dos alunos nos exames. A intenção deste projeto foi criar mecanismos de incentivo de esforço dos professores e de atrair e manter bons profissionais para a área de educação. Os resultados mostraram melhora de cerca de 40% nas notas.

Como visto, boa parte da literatura internacional mostra indícios de impactos positivos sobre as notas associados aos sistemas de incentivos aos professores por meio de premiação. Este artigo pretende contribuir para o estudo na área por meio de um trabalho empírico a respeito do efeito de premiação de docentes sobre o desempenho dos alunos no Brasil.

3 – O programa de bonificação de São Paulo

Para a análise do impacto de bonificação de professores sobre proficiência dos alunos no estado de São Paulo, é importante detalhar os principais aspectos do programa que chegou a pagar em 2010 cerca de R\$ 650 milhões em bônus a aproximadamente 210 mil funcionários da rede estadual.

As regras do sistema de remuneração foram lançadas pelo Programa de Qualidade da Escola (PQE) em maio de 2008 e o projeto de lei referente ao programa foi aprovado apenas em dezembro do mesmo ano⁵. Em março de 2009, o governo estadual remunerou cerca de 196 mil servidores baseado em resultados de 2008.

A Secretaria de Educação do Estado de São Paulo estipula metas para a série final de cada ciclo de ensino - quarta série (quinto ano) e oitava série (nono ano) do ensino fundamental e terceiro ano do ensino médio - da escola e remunera todos os professores do ciclo avaliado e os demais funcionários. O desenho do programa pode ser interessante no sentido de que toda a equipe pode cooperar para um melhor ambiente escolar. Segundo Lavy (2009), em educação, incentivos para grupos são melhores que incentivos individuais pela própria natureza do processo educacional, que envolve trabalho em equipe, esforços de colegas professores e multitarefas. Embora não seja possível descartar a existência de *free-riders*, as regras do programa tentam eliminar o problema do absenteísmo já que funcionários com mais de 1/3 de faltas não têm direito ao bônus.

As metas são estabelecidas com base no Índice de Desenvolvimento da Educação do Estado de São Paulo (IDESP) para cada etapa escolar (s), que por sua vez, é composto por dois outros índices (Secretaria da Educação do Estado de São Paulo, 2009):

$$IDESP_s = ID_s \times IF_s \quad (1)$$

No qual ID é o índice de Desempenho e IF é o índice de Fluxo e serão detalhados mais a frente. A ideia por trás dessa composição é que uma escola pode ser considerada boa quando a maior parte dos alunos desenvolve as competências e habilidades requeridas para a sua série num período de tempo ideal - o ano letivo.

⁵ Projeto de Lei Complementar no. 1078

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. O exame 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. São aplicadas provas de Língua Portuguesa com Redação, Matemática e Ciências Humanas (Geografia e História). Para o pagamento de bônus, considera-se apenas a proficiência em Matemática e Português.

O ID representa a quantidade relativa de alunos em cada nível de proficiência. São quatro níveis: (i) Abaixo do básico, (ii) Básico, (iii) Adequado e (iv) Avançado. Os valores de referência do SARESP para a definição dos níveis de desempenho encontram-se na tabela a seguir:

Tabela 1: Valores de referência para níveis de proficiência

	Quarta série		Oitava série	
	Português	Matemática	Português	Matemática
Abaixo do básico	<150	<175	<200	<225
Básico	entre 150 e 200	entre 175 e 225	entre 200 e 275	entre 225 e 300
Adequado	entre 200 e 250	entre 225 e 275	entre 275 e 325	entre 300 e 350
Avançado	>250	> 275	> 325	> 350

Fonte: Secretaria de Educação do Estado de São Paulo

Assim, o ID para cada disciplina (j) e para cada etapa de escolarização (s) é dado por:

$$ID_{js} = \left(1 - \frac{def_{js}}{3}\right) \times 10 \quad (2)$$

Na fórmula, def é um indicador que mede a defasagem em relação às expectativas de aprendizado.

$$def_{js} = \frac{(3 \times AB) + (2 \times B) + (1 \times Ad) + (0 \times Av)}{100} \quad (3)$$

Em que AB , B , Ad e Av representam respectivamente a proporção de alunos nos níveis abaixo do básico, básico, adequado e avançado em relação ao total de alunos avaliados.

Quando a escola apresenta grande defasagem, ou seja, nível de proficiência baixo, o índice de desempenho é baixo. O ID varia de zero a dez e a média aritmética entre os índices de Língua Portuguesa (LP) e Matemática (Mat) representa o ID para cada etapa de escolarização.

$$ID_s = \frac{ID_{LP} + ID_{Mat}}{2} \quad (4)$$

O indicador de Fluxo (IF) é dado pela taxa média de aprovação em cada etapa de escolarização e é dado por:

$$IF_s = \frac{A_i + A_{i+1} + \dots + A_n}{n} = \frac{\sum_{i=1}^n A_i}{n} \quad (5)$$

Em que A_i é a taxa de aprovação na série i e n é o número de séries⁶ da etapa de escolarização considerada.

$$A_i = \frac{\text{número de alunos aprovados}}{\text{número total de alunos matriculados}} \quad (6)$$

⁶ Atualmente, com o ensino fundamental de nove anos, n é igual a 5 para as séries iniciais do EF e 4 para as séries finais.

A partir dos dados do IDESP calcula-se a parcela que a escola cumpriu da meta. O Índice de Cumprimento (IC) em determinado ano (t) para cada etapa escolar (s) é dado pela seguinte fórmula:

$$IC_{s,t} = \left(\frac{IDESP_{s,t} - IDESP_{s,t-1}}{META_{s,t} - IDESP_{s,t-1}} \right) \quad (7)$$

O IC é limitado no intervalo de 0% a 120%. Para 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 (META 2030). O Índice de Qualidade (IQ) do ano t para a série s é calculado pela seguinte fórmula:

$$IQ_{s,t} = \left(\frac{IDESP_{s,t} - IDESP(global)_{s,t}}{META_{s,2030} - IDESP(global)_{s,t}} \right) \quad (8)$$

Para efeito de pagamento, considera-se o indicador (IQ+IC), que está limitado entre 0 e 120%.

As metas de longo prazo do IDESP para cada série são iguais para todas as escolas e foram estabelecidas levando-se em consideração níveis de aprendizado comparáveis aos de países da Organização para a Cooperação e o Desenvolvimento Econômico (OCDE). As metas para 2030 para quarta e oitava séries do fundamental são, respectivamente, 7,0 e 6,0.

Em 2009, o IDESP médio para a quarta série da rede estadual foi de 3,86. Já para a oitava série foi de 2,84. As escolas têm, portanto, um grande desafio até 2030, pois, na média, estão muito abaixo do nível desejado.

Visando atingir este nível de pontuação, foram estabelecidas metas anuais (que entram no cálculo do IC e IQ) para cada escola de acordo com 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 relativamente àquelas com maiores notas. Embora a variação seja maior, é razoável acreditar que em escolas piores, menores esforços possam gerar maiores mudanças de notas. Portanto, embora escolas melhores tenham que apresentar uma menor variação de notas, como já se encontram em um patamar elevado, os esforços necessários podem ser maiores e o cumprimento da meta pode ser mais difícil. Nesse sentido, torna-se interessante verificar se o desenho do programa é capaz de gerar incentivos tanto para escolas mais fracas quanto para escolas mais fortes, uma vez que elas devem ser muito diferentes em vários aspectos.

Com relação ao pagamento do bônus, o valor varia de acordo com a parcela da meta cumprida e é feito a professores do ciclo que cumpriu 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. Escolas que não cumprem as metas, mas que nos dois anos anteriores foram classificadas entre as 10% melhores, também recebem uma remuneração. Este seria um incentivo para escolas boas, cujas metas são difíceis de cumprir, pois as notas já estão num patamar elevado. Funcionários das diretorias regionais recebem bônus pela média da região em que atuam.

4 – Metodologia

Neste trabalho, pretende-se avaliar o efeito da política de *accountability* forte adotada no estado de São Paulo a partir de 2008 sobre o desempenho dos alunos. Para averiguar o impacto deste programa, deveríamos comparar o desempenho em duas situações, com e sem a presença da política. No entanto, é impossível observar essas duas situações simultaneamente.

Segundo Abadie et al. (2001), 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 da *accountability*.

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 na forma de bônus em 2008.

Para lidar com esse problema, a metodologia adotada foi o *matching*, ou pareamento. A abordagem consiste em separar os indivíduos em dois grupos (tratamento e controle), considerando características observáveis. Deste modo, a aleatoriedade da escolha estaria garantida, condicional em características observáveis. Em nosso trabalho, o grupo de tratamento seria formado por escolas estaduais paulistas, enquanto o de controle abrange escolas de estados que não contam com esse programa e escolas das redes municipais do estado de São Paulo.

Após o pareamento, podemos eliminar os efeitos de características não observáveis que afetam o desempenho e que são invariantes no tempo. Fazemos isso, utilizando um estimador de Diferenças em Diferenças e comparando os resultados nas duas situações.

Para formalizar o procedimento que será realizado, considere o status tratamento de um indivíduo i por meio de uma variável binária D que assume o valor 1 quando participa do programa e 0, caso contrário. Em adição, suponha que o valor da variável de interesse para o indivíduo i seja Y_{i1} , caso ele seja submetido ao tratamento (1) e Y_{i0} , caso não seja exposto ao tratamento (0). O efeito do tratamento para o indivíduo i (τ_i) é dado por $\tau_i = Y_{i1} - Y_{i0}$

O efeito médio do tratamento sobre os tratados (*average treatment effect on treated* – ATT) é dado por:

$$\tau_{ATT} = E(Y_{i1} - Y_{i0}|D_i = 1) = E(Y_{i1}|D_i = 1) - E(Y_{i0}|D_i = 1) \quad (9)$$

No entanto, o contrafactual de um indivíduo sob tratamento, $E(Y_{i0}|D_i = 1)$, não é observável.

O efeito médio do tratamento sobre os não tratados (*average treatment effect on the control* – ATC) é dado por:

$$\tau_{ATC} = E(Y_{i1} - Y_{i0}|D_i = 0) = E(Y_{i1}|D_i = 0) - E(Y_{i0}|D_i = 0) \quad (10)$$

Neste caso, o contrafactual do indivíduo que não recebe tratamento é $E(Y_{i1}|D_i = 0)$, que não é observável.

Os resultados observáveis são, portanto, apenas $E(Y_{i1}|D_i = 1)$ e $E(Y_{i0}|D_i = 0)$. Se pudéssemos considerar que $E(Y_{i0}|D_i = 1) = E(Y_{i0}|D_i = 0)$, poderíamos substituir $E(Y_{i0}|D_i = 0)$ na equação (9) e então teríamos o ATT. Mas, essa é uma suposição muito forte, pois indivíduos tratados e não tratados são diferentes. Uma estimação feita desta forma muito provavelmente estaria viesada.

É possível evitar esse viés se for assumida a Hipótese da Independência Condicional (*Conditional Independence Assumption* - CIA), que garante que a escolha do grupo de tratamento é independente dos resultados. Além disso, é necessário garantir que, dado um conjunto de características observáveis X , indivíduos com as mesmas características têm a mesma probabilidade de participar do grupo tratamento ou controle. Dessa forma:

$$Y_{i1}, Y_{i0} \perp D_i | X_i \quad (11)$$

$$Y_{i1}, Y_{i0} \perp D_i | X_i \Rightarrow (Y_{i0}|X_i, D_i = 1) = (Y_{i0}|X_i, D_i = 0) = (Y_{i0}|X_i) \quad (12)$$

em que \perp denota independência.

A CIA afirma que os resultados potenciais independem da participação do programa dadas as características observáveis X .

Podemos reescrever a equação (9) da seguinte forma:

$$\tau_{ATT} = E(Y_{i1} - Y_{i0}|D_i = 1, X) = E(Y_{i1}|D_i = 1, X) - E(Y_{i0}|D_i = 0, X) \quad (13)$$

Assim, se vale a CIA, teremos uma estimativa não viesada do ATT. Utilizaremos a metodologia de *matching* (ou pareamento) que consiste em encontrar um grupo de controle o mais parecido possível ao grupo de tratamento de modo que os resultados desses possam ser comparados. Assim, condicionando nas variáveis X , é possível comparar os dois grupos como se fosse um experimento aleatório.

Quanto maior o número de características observáveis, maior a semelhança entre os indivíduos, tornando-os comparáveis. Mas, se o número de variáveis observadas for muito grande, menor é a probabilidade de encontrar alguém parecido. Ao considerar todas as covariadas relevantes, o pareamento pode apresentar o problema da multidimensionalidade. Para n características observáveis, o número de possíveis valores para o vetor X será 2^n .

Um modo de resolver o problema da multidimensionalidade é o *propensity score* (escore de propensão), proposto por Rosenbaun e Rubin (1983). O método consiste em estimar uma probabilidade, $P(X)$, do indivíduo receber tratamento, dadas suas covariadas. Desta forma, o pareamento é feito com base em apenas um escalar.

Esse método atribuirá uma probabilidade entre zero e um da escola receber tratamento. Seja $P(X)$ a probabilidade de um indivíduo i receber tratamento, definida como:

$$P(X) = Pr(D_i = 1|X) \quad (14)$$

Segundo esta metodologia, X pode ser substituído por $P(X)$:

$$\begin{aligned} Y_{i1}, Y_{i0} &\perp D_i | X_i \\ Y_{i1}, Y_{i0} &\perp D_i | P(X_i) \end{aligned}$$

Mesmo controlando pelas variáveis observadas, ainda podem existir características não observadas que afetam o resultado de interesse. Para este trabalho, a base de dados nos permite utilizar o método de Diferenças em Diferenças (DD) para eliminar o efeito de características individuais não observáveis que são fixas no tempo.

O DD consiste na comparação da diferença dos resultados (Y) do grupo de tratamento e controle ao longo do tempo (diferença no tempo). Considere Y^A , resultado antes da intervenção Y^D o resultado depois da intervenção, DD determina o efeito da intervenção sobre o resultado de interesse Y . Dito de outra forma:

$$DD = E(Y_1^D - Y_1^A | D = 1) - E(Y_0^D - Y_0^A | D = 1) \quad (15)$$

No entanto a última parcela do lado direito da equação $E(Y_0^D - Y_0^A | D = 1)$ não é observável. Assim, para que seja possível calcular o estimador de DD, é necessário admitir a Hipótese de Tendência Comum que é dada por: $E(Y_0^D - Y_0^A | D = 1) = E(Y_0^D - Y_0^A | D = 0)$.

Se a hipótese é válida, o impacto médio do programa é dado por:

$$DD = E(Y_1^D - Y_1^A | D = 1) - E(Y_0^D - Y_0^A | D = 0) \quad (16)$$

Uma vez que já foi determinado o grupo de comparação por meio do *matching*, o DD pode ser escrito como a diferença do ATT antes e depois da intervenção.

$$DD = [E(Y_1^D | D = 1) - E(Y_0^D | D = 0)] - [E(Y_1^A | D = 1) - E(Y_0^A | D = 0)] \quad (17)$$

$$\beta_1 = \tau_{ATT}^D - \tau_{ATT}^A \quad (18)$$

O estimador DD também pode ser calculado pelo método de Mínimos Quadrados Ordinários a partir do seguinte modelo:

$$Y_{it} = \beta_0 + \beta_1 D_i t + \beta_2 D_i + \gamma t + \varepsilon_{it} \quad (19)$$

Em que:

$D_i = 1$, se recebeu tratamento e 0, caso contrário

$t = 1$, depois do tratamento e 0, antes.

No nosso caso, o β_1 estimado será o impacto do programa sobre o desempenho do aluno.

Nossa análise em termos das notas será feita com base em prova de abrangência nacional, a Prova Brasil. Em outras palavras, a intenção é verificar se as notas da Prova Brasil em escolas paulistas

estaduais tiveram efeito diferenciado quando comparadas com as notas das escolas municipais paulistas e com as escolas de outros estados.

Essa análise é possível, uma vez que entendemos que a política de *accountability* forte por meio da bonificação a partir do desempenho do SARESP tem também impacto positivo no desempenho da Prova Brasil.

É plausível supor que haja uma correlação forte entre os dois exames, já que os testes padronizados utilizam a mesma metodologia em provas de Matemática e Língua Portuguesa desde 2007.

5 – Base de dados

Nossa variável de interesse neste trabalho é a proficiência em Matemática e Português dos alunos de quarta e oitava séries apurada pela Prova Brasil. Os questionários aplicados à alunos, professores e diretores que realizam este exame fornecem as covariadas necessárias para a estimação do *Propensity Score Matching*.

O quadro abaixo mostra as descrições dos dados da Prova Brasil de 2007. As características de escolas e diretores são variáveis *dummy*, já para as variáveis de professores e alunos consideraram-se as proporções destes dentro da escola.

Tabela 2: Descrição das variáveis

Descrição das variáveis
Variáveis de alunos
Proporção de alunos pardos
Proporção de alunos com computador e Internet
Proporção de alunos que moram com 1 pessoa
Proporção de alunos cuja mãe completou a oitava, mas não completou Ens. Médio
Proporção de alunos cujos pais sempre vão à reunião
Proporção de alunos que possuem de 1 a 20 livros em casa
Proporção de alunos que possuem de 21 a 100 livros em casa
Proporção de alunos que possuem mais de 100 livros em casa
Proporção de alunos que trabalham fora
Proporção de alunos que fizeram o maternal
Proporção de alunos de fizeram a pré-escola
Proporção de alunos que nunca reprovaram
Proporção de alunos que moram com os pais
Proporção de alunos que moram com pai ou mãe
Variáveis de professor
Proporção de professor com até 2 anos de experiência
Proporção de professor de 3 a 9 anos de experiência
Proporção de professor com salário de até R\$1300
Variáveis de diretor
Variável <i>dummy</i> : Salário Bruto do diretor: mais de R\$1500
Variável <i>dummy</i> : Experiência do diretor - até 4 anos
Variável <i>dummy</i> : Experiência do diretor - 5 a 15 anos
Variável <i>dummy</i> : Conselho escolar - tem e se reúne pelo menos uma vez/ano
Variável <i>dummy</i> : Conselho de classe - tem e se reúne pelo menos uma vez/ano

Fonte: Ministério da Educação – Prova Brasil 2007

6 – Resultados

Estimamos o escore de propensão por meio de modelo Probit. No contexto deste trabalho, será estimada a probabilidade de ser uma escola estadual paulista. A Prova Brasil 2007 nos permite utilizar um número grande de regressores. Quanto mais covariadas são usadas, mais o grupo de controle ficará

parecido com o grupo de tratamento, porém, a dificuldade de encontrar um controle aumenta. As especificações que apresentaram uma boa qualidade do *matching*, com covariadas balanceadas e maior região de suporte comum foram as seguintes:

Tabela 3: Probit para Quarta e Oitava série – Matemática e Português

Variável dependente:	Quarta série		Oitava série	
	Matemática	Português	Matemática	Português
Escola estadual paulista				
Variáveis de alunos				
Pardo	0.705*** (0.133)	0.730*** (0.133)	-	-
Tem computador com internet	2.240*** (0.160)	2.285*** (0.160)	0.897*** (0.0932)	0.915*** (0.0959)
Escolaridade da mãe (8 a 10 anos)	1.654*** (0.312)	1.665*** (0.311)	1.673*** (0.226)	1.637*** (0.229)
Possui de 1 a 20 livros em casa	-2.513*** (0.214)	-2.489*** (0.214)	-2.139*** (0.227)	-2.110*** (0.230)
Possui de 21 a 100 livros em casa	-3.999*** (0.294)	-4.032*** (0.294)	-3.530*** (0.239)	-3.478*** (0.244)
Possui mais de 100 livros em casa	-3.458*** (0.365)	-3.492*** (0.365)	-3.651*** (0.391)	-3.404*** (0.398)
Trabalha fora	4.258*** (0.265)	4.322*** (0.266)	-	0.445*** (0.163)
Fez maternal	-1.422*** (0.115)	-1.412*** (0.116)	-0.599*** (0.0814)	-0.616*** (0.0828)
Nunca foi reprovado	6.341*** (0.198)	6.333*** (0.198)	6.717*** (0.143)	6.967*** (0.147)
Mora com os pais	1.080*** (0.386)	1.079*** (0.384)	2.745*** (0.323)	3.067*** (0.329)
Mora com o pai ou a mãe	1.607*** (0.433)	1.578*** (0.431)	5.862*** (0.353)	6.372*** (0.361)
Variáveis de professores				
Experiência de até 2 anos	0.742*** (0.0528)	0.774*** (0.0535)	-	-
Experiência de 3 a 9 anos	-0.370*** (0.0614)	-0.334*** (0.0621)	-	-
Salário até R\$1300	-1.115*** (0.0773)	-1.083*** (0.0756)	-0.751*** (0.0500)	-0.999*** (0.0533)
Variáveis de diretores				
Salário mais de R\$1500	0.827*** (0.0650)	0.816*** (0.0649)	1.019*** (0.0571)	1.006*** (0.0575)
Experiência de até 4 anos	-0.682*** (0.0526)	-0.682*** (0.0526)	-	-
Experiência de 5 a 15 anos	-0.201*** (0.0488)	-0.206*** (0.0488)	-	-
Conselho de escola se reúne pelo menos uma vez/ano	1.029*** (0.101)	1.034*** (0.101)	-	-
Conselho de classe se reúne pelo menos uma vez/ano	0.547*** (0.115)	0.563*** (0.116)	-	-
Constante	-8.065*** (0.417)	-8.125*** (0.415)	-8.147*** (0.311)	-8.765*** (0.327)
No. de observações	32,064	32,083	23,201	23,204
Pseudo R-quadrado	0.468	0.469	0.408	0.421
Erro padrão entre parênteses	*** p<0.01, ** p<0.05, * p<0.1			

Fonte: Elaboração própria

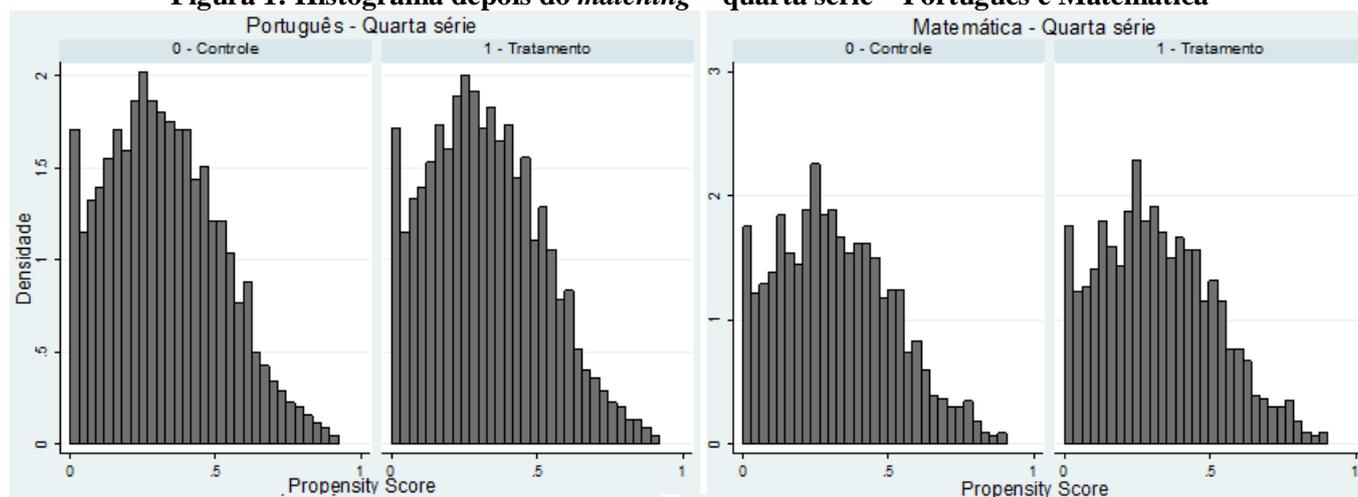
Após a estimação do escore de propensão foi feito o *matching* utilizando o método do primeiro vizinho mais próximo. A tabela 4 mostra que o número de tratados e controles dentro do suporte comum é relativamente grande. A boa aderência do *matching* também pode ser notada ao observar a distribuição do escore de propensão nos dois grupos nas figuras 1 e 2.

Tabela 4: Suporte comum

		Suporte Comum			
			Fora do suporte	Dentro do suporte	Total
Quarta série	Matemática	Controles	0	30311	30311
		Tratados	266	1487	1753
	Português	Controles	0	30329	30329
		Tratados	260	1494	1754
Oitava série	Matemática	Controles	0	19871	19871
		Tratados	720	2610	3330
	Português	Controles	0	19865	19865
		Tratados	764	2575	3339

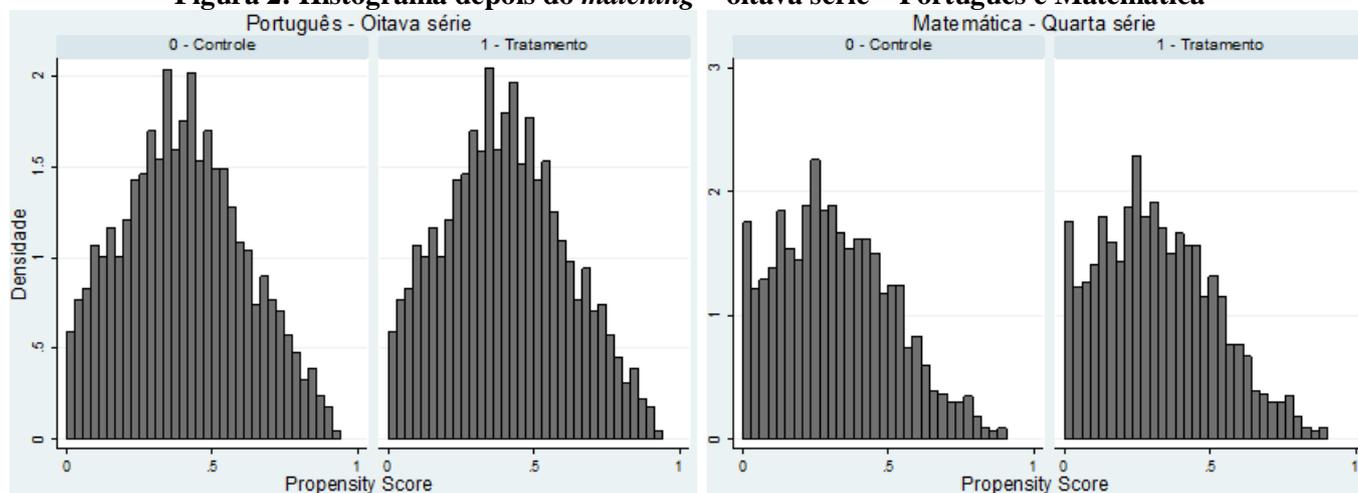
Fonte: Elaboração própria

Figura 1: Histograma depois do *matching* – quarta série – Português e Matemática



Fonte: Elaboração própria

Figura 2: Histograma depois do *matching* – oitava série – Português e Matemática



Fonte: Elaboração própria

Outro procedimento importante é a checagem das condições de balanceamento. As tabelas A1 a A4 do Apêndice mostram a média das variáveis no tratamento e no controle. A hipótese nula do teste é que a diferença das médias é zero, ou seja, as médias dos dois grupos são iguais. Antes do pareamento, é esperado que as médias fossem diferentes. Depois de pareados, percebe-se que em quase todos os casos tem-se um *matching* com um bom balanceamento de covariadas.

Por fim, pode-se também checar o que ocorreu com o Pseudo R². Ao se reestimar o escore de propensão utilizando apenas informações das escolas tratadas e de controle selecionadas o valor do Pseudo R-quadrado deve diminuir, se o *matching* tiver sido bem realizado. De fato, foi o que ocorreu, como mostrado na Tabela 5. Após o pareamento, 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.

Tabela 5: Pseudo R²

	Pseudo- R ²			
	Quarta série		Oitava série	
	Matemática	Português	Matemática	Português
Antes do Matching	0,468	0,469	0,408	0,421
Depois do Matching	0,004	0,002	0,001	0,001

Fonte: Elaboração própria

Embora os procedimentos anteriores indiquem que o pareamento foi bem feito, as notas médias entre grupos de controle e tratamento no período pré-tratamento (2007) ainda mostraram-se estatisticamente diferentes para quarta e oitava séries (Tabela 6). Essa diferença indica que características não observáveis podem estar afetando a proficiência dos alunos e que devem ser consideradas através do método de diferenças em diferenças.

Tabela 6: Matching por vizinho mais próximo (quarta e oitava série)

		Amostra	Tratados	Controles	Diferença	D. Padrão	Estatística t
Quarta série	Matemática	Antes do Matching	193,70	189,21	4,49	0,49	9,15
		Depois do Matching	193,01	203,29	-10,28	0,64	-16,10
	Português	Antes do Matching	176,65	171,73	4,92	0,45	11,02
		Depois do Matching	176,19	184,42	-8,24	0,59	-14,06
Oitava série	Matemática	Antes do Matching	242,13	239,04	3,09	0,35	8,77
		Depois do Matching	241,20	250,73	-9,53	0,45	-21,15
	Português	Antes do Matching	231,26	227,16	4,10	0,32	12,96
		Depois do Matching	230,13	237,26	-7,13	0,41	-17,40

Obs: Nearest Neighbor (1) com Caliper (0.002), sem reposição.

Fonte: Elaboração própria

Após o pareamento, o próximo passo é aplicar o estimador de Diferenças em Diferenças que permite eliminar o efeito de características não observadas (fixas no tempo) que podem afetar o resultado.

Utilizando o método de MQO, estimou-se o seguinte modelo:

$$Nota\ média_{it} = \beta_0 + \beta_1 Ano * Tratamento_i + \beta_2 Tratamento_i + \gamma Ano + \varepsilon_{it} \quad (20)$$

Em que *Ano* é igual a 1 para 2009 e 0 para 2007. Desta forma, o valor de β_1 representará o impacto do programa sobre o desempenho do aluno tratado.

Os resultados (Tabela 7) mostram indícios de melhora da nota dos alunos com a implantação do programa na rede estadual paulista. Entretanto, houve diferenças nos efeitos entre disciplinas e ciclo escolar. Os impactos foram positivos apenas para a quarta série. O efeito estimado para Matemática foi bastante forte, de cerca de 6,4 pontos na escala SAEB. O efeito em Português foi um pouco menor, mas

ainda relevante e significativo, de 3,7 pontos. Com relação à oitava série os coeficientes de interesse não foram significativos estatisticamente.

Tabela 7: Resultado do Diferenças em Diferenças

Variáveis	Quarta série		Oitava série	
	Matemática	Português	Matemática	Português
Tratamento	-10.28*** (0.678)	-8.236*** (0.596)	-9.526*** -0,468	-7.127*** -0,43
Ano	13.60*** (0.678)	9.263*** (0.596)	0.906* (0.468)	9.492*** (0.430)
Ano*Tratamento	6.357** (0.959)	3.727*** (0.843)	0,0554 (0.662)	-0,504 (0.609)
Constante	203.3*** (0.479)	184.4*** (0.422)	250.7*** (0.331)	237.3*** (0.304)
Observações	5948	5976	10440	10300
R-quadrado	0,2	0,137	0,074	0,128

Erro Padrão entre parênteses

*** p<0.01, ** p<0.05, * p<0.1

Fonte: Elaboração própria

De modo geral, os resultados indicam um efeito de curto prazo positivo da política de bonificação. É possível que efeitos diferentes sejam encontrados ao considerar os dados dos próximos exames, passado o período inicial de adaptação das escolas e do corpo docente a essas formas de avaliação com expectativa de premiação.

7 – Conclusão

Dentro do contexto de responsabilização dos agentes escolares pelos resultados dos alunos, este trabalho procurou investigar se o programa de bonificação de professores implementado no final de 2008 no estado de São Paulo teve impacto sobre a proficiência dos estudantes.

Utilizou-se o método de *Propensity Score Matching* (PSM) para buscar grupos comparáveis e depois uma estimação de Diferenças em Diferenças (DD) para eliminar efeitos fixos não observáveis. Os procedimentos de checagem mostraram boa aderência do *matching*.

Estudos em vários países indicam que programas de premiação de docentes levam, em geral, a uma melhora do desempenho dos alunos. Os resultados encontrados neste trabalho mostraram-se, parcialmente, em linha com os apresentados pela literatura internacional.

Foram encontrados indícios de impacto positivo do programa de incentivo de professores apenas nas notas da quarta série. O efeito para Matemática foi de cerca de 6,4 pontos e o de Língua Portuguesa em cerca de 3,7 pontos. Para a oitava série não houve evidência de efeito significativo.

Referências Bibliográficas

ABADIE,A.,DIAMOND,A.;HAINMUELLER,J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. **National Bureau of Economic Research**. v. 12831, 2007.

_____, DRUKKER, D; IMBENS, G. Implementing matching estimators for average treatment effects in Stata. **The Stata Journal**. College Station, v.4, n.3, p.290-311, 2001.

ATKINSON, A.; BURGESS,S.; CROXSON,B; GREGG,P.;PROPPER,C.;SLATER, H.;WILSON,D. Evaluating the impact of performance-related pay for teachers in England. **Labour Economics** v.16, n.3, p.251–26, 2009.

DOLTON,P.; MARCENARO-GUTIERREZ,O.D. If you pay peanuts do you get monkeys? A cross country analysis of teacher pay and pupil performance, **Economic Policy** January 2011 pp. 5–55, 2001.

FERNANDES,R; GREMAUD,A. Qualidade da educação: avaliação, indicadores e metas. In: VELOSO,F. (Org), **Educação Básica no Brasil**. Rio de Janeiro: Elsevier, 2009.

FERRAZ,C. Sistemas Educacionais Baseados em Desempenho, Metas de Qualidade e a Remuneração de Professores: os Casos de Pernambuco e São Paulo. In: VELOSO,F. (Org), **Educação Básica no Brasil**. Rio de Janeiro: Elsevier, 2009.

FIGLIO,D.N.; KENNY, L.W. Individual teacher incentives and student performance. *Journal of Public Economics*,v.91, p.901–914, 2007.

GLEWWE,P.; ILIAS,N.;KREMER,M. Teacher Incentives. National Bureau of Economic Research, v.9671, 2003.

HANUSHEK,E. The failure of input-based schooling policies. *Economic Journal*, v.113, p.64-98, 2003.
HOLMSTROM,B.;MILGROM,P. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization*, v. 7, p. 24-52, 1991.

LAVY,V. Performance Pay and Teachers' Effort, Productivity, and Grading Ethics. **American Economic Review**, v.99.n.5.p.1979–2011, 2009

MINISTÉRIO DA EDUCAÇÃO - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **SAEB – 2005 – Primeiros Resultados: Médias de desempenho do SAEB/2005 em perspectiva comparada**. 2007 disponível em: http://www.inep.gov.br/download/saeb/2005/SAEB1995_2005.pdf

ROSENBAUM, P.R., RUBIN, D.B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, Great Britain, v.70, n.1, p.41-55, 1983.

SÃO PAULO. Lei Complementar 1078 de 17 de dezembro de 2008. Institui Bonificação por Resultados - BR, no âmbito da Secretaria da Educação, e dá providências correlatas. **Diário Oficial [do] Estado de São Paulo** no. 239, volume 118 de 18 de dezembro de 2008.

SECRETARIA DE EDUCAÇÃO DE SÃO PAULO. **Programa de Qualidade da Escola**. 2009. Nota Técnica disponível em: <http://idesp.edunet.sp.gov.br/Arquivos/NotaTecnicaPQE2008.pdf>

_____. **Programa de Qualidade da Escola**. 2010. Nota Técnica disponível em: http://idesp.edunet.sp.gov.br/Arquivos/Nota_tecnica_2009.pdf

TODOS PELA EDUCAÇÃO. **Números do Brasil**. Disponível em: <http://www.todospelaeducacao.org.br/educacao-no-brasil/numeros-do-brasil/brasil>. Acesso em: 08 março 2011.

VELOSO, F. Experiências de Reforma Educacional nas Últimas Duas Décadas: o que Podemos Aprender?. In: VELOSO, F. (Org), **Educação Básica no Brasil**. Rio de Janeiro: Elsevier, 2009.

VIGDOR, J.L. **Teacher Salary Bonuses in North Carolina**. National Center for analysis of longitudinal data in education research. Working Paper n.15, 2008.

APÊNDICE

Tabela A1: Diferença das médias das variáveis explicativas – Quarta série – Matemática

Matemática - quarta série	Antes do Matching			Depois do Matching		
	Tratamento	Controle	p-valor	Tratamento	Controle	p-valor
Variáveis de alunos						
Pardo	0,422	0,415	0,076	0,421	0,419	0,62
Tem computador com internet	0,291	0,158	0	0,286	0,290	0,45
Escolaridade da mãe (8 a 10 anos)	0,141	0,125	0	0,141	0,142	0,54
Possui de 1 a 20 livros em casa	0,523	0,520	0,255	0,523	0,521	0,36
Possui de 21 a 100 livros em casa	0,156	0,157	0,564	0,157	0,158	0,78
Possui mais de 100 livros em casa	0,091	0,091	0,525	0,091	0,090	0,69
Trabalha fora	0,129	0,135	0,003	0,128	0,125	0,15
Fez maternal	0,339	0,334	0,255	0,341	0,336	0,43
Nunca foi reprovado	0,813	0,590	0	0,803	0,806	0,47
Mora com os pais	0,618	0,532	0	0,617	0,622	0,14
Mora com o pai ou a mãe	0,324	0,335	0	0,323	0,319	0,12
Variáveis de professores						
Experiência de até 2 anos	0,541	0,311	0	0,515	0,506	0,49
Experiência de 3 a 9 anos	0,223	0,397	0	0,242	0,257	0,17
Salário até R\$1300	0,078	0,317	0	0,085	0,079	0,31
Variáveis de diretores						
Salário mais de R\$1500	0,959	0,566	0	0,953	0,952	0,86
Experiência de até 4 anos	0,251	0,552	0	0,278	0,295	0,29
Experiência de 5 a 15 anos	0,558	0,375	0	0,553	0,532	0,25
Conselho de escola se reúne pelo menos uma vez/ano	0,987	0,823	0,000	0,985	0,978	0,221
Conselho de classe se reúne pelo menos uma vez/ano	0,986	0,797	0,000	0,984	0,985	0,883

Fonte: Elaboração própria

Tabela A2: Diferença das médias das variáveis explicativas – Quarta série – Português

Português - quarta série	Antes do Matching			Depois do Matching		
	Tratamento	Controle	p-valor	Tratamento	Controle	p-valor
Variáveis de alunos						
Pardo	0,422	0,415	0,076	0,421	0,420	0,821
Tem computador com internet	0,292	0,158	0	0,287	0,291	0,530
Escolaridade da mãe (8 a 10 anos)	0,141	0,125	0	0,141	0,141	0,950
Possui de 1 a 20 livros em casa	0,523	0,520	0,259	0,524	0,521	0,289
Possui de 21 a 100 livros em casa	0,156	0,157	0,558	0,157	0,158	0,849
Possui mais de 100 livros em casa	0,091	0,091	0,572	0,092	0,092	0,702
Trabalha fora	0,129	0,135	0,003	0,127	0,127	0,971
Fez maternal	0,339	0,334	0,231	0,341	0,342	0,916
Nunca foi reprovado	0,813	0,590	0	0,805	0,805	0,999
Mora com os pais	0,618	0,532	0	0,618	0,617	0,849
Mora com o pai ou a mãe	0,324	0,335	0	0,323	0,322	0,749
Variáveis de professores						
Experiência de até 2 anos	0,552	0,315	0	0,527	0,524	0,837
Experiência de 3 a 9 anos	0,226	0,398	0	0,241	0,252	0,295
Salário até R\$1300	0,082	0,317	0	0,089	0,081	0,200
Variáveis de diretores						
Salário mais de R\$1500	0,959	0,566	0	0,954	0,950	0,608
Experiência de até 4 anos	0,251	0,552	0	0,277	0,280	0,838
Experiência de 5 a 15 anos	0,558	0,375	0	0,557	0,542	0,398
Conselho de escola se reúne pelo menos uma vez/ano	0,987	0,824	0	0,985	0,977	0,141
Conselho de classe se reúne pelo menos uma vez/ano	0,986	0,797	0	0,985	0,986	0,878

Fonte: Elaboração própria

Tabela A3: Diferença das médias das variáveis explicativas – Oitava série – Português

Português - oitava	Antes do Matching			Depois do Matching		
	Tratamento	Controle	p-valor	Tratamento	Controle	p-valor
Variáveis de alunos						
Tem computador com internet	0,311	0,181	0,0	0,302	0,310	0,120
Escolaridade da mãe (8 a 10 anos)	0,184	0,152	0,0	0,181	0,183	0,438
Possui de 1 a 20 livros em casa	0,603	0,592	0,0	0,601	0,601	0,989
Possui de 21 a 100 livros em casa	0,183	0,193	0,0	0,185	0,187	0,391
Possui mais de 100 livros em casa	0,055	0,064	0,0	0,057	0,056	0,392
Trabalha fora	0,182	0,211	0,0	0,184	0,184	0,981
Fez maternal	0,422	0,418	0,4	0,421	0,421	0,968
Nunca foi reprovado	0,787	0,593	0,0	0,769	0,769	0,796
Mora com os pais	0,618	0,596	0,0	0,618	0,620	0,505
Mora com o pai ou a mãe	0,321	0,292	0,0	0,318	0,317	0,662
Variáveis de professores						
Salário até R\$1300	0,066	0,278	0,0	0,078	0,086	0,222
Variáveis de professores						
Salário mais de R\$1500	0,972	0,661	0,0	0,964	0,964	0,881

Fonte: Elaboração própria

Tabela A4: Diferença das médias das variáveis explicativas – Oitava série – Matemática

Matemática - oitava	Antes do Matching			Depois do Matching		
	Tratamento	Controle	p-valor	Tratamento	Controle	p-valor
Variáveis de alunos						
Tem computador com internet	0,311	0,182	0	0,302	0,308	0,224
Escolaridade da mãe (8 a 10 anos)	0,184	0,152	0	0,181	0,182	0,594
Possui de 1 a 20 livros em casa	0,603	0,592	0	0,602	0,600	0,418
Possui de 21 a 100 livros em casa	0,183	0,193	0	0,185	0,188	0,194
Possui mais de 100 livros em casa	0,055	0,064	0	0,056	0,056	0,745
Fez maternal	0,422	0,418	0,368	0,421	0,418	0,559
Nunca foi reprovado	0,786	0,593	0	0,768	0,769	0,752
Mora com os pais	0,618	0,596	0	0,617	0,621	0,099
Mora com o pai ou a mãe	0,321	0,291	0	0,319	0,316	0,172
Variáveis de professores						
Salário até R\$1300	0,074	0,279	0	0,086	0,085	0,846
Variáveis de professores						
Salário mais de R\$1500	0,972	0,661	0	0,965	0,964	0,823

Fonte: Elaboração própria