

Avaliando o Impacto do Caderno de Apoio e Aprendizagem na Rede de Ensino Municipal de São Paulo: Efeitos Médios e Heterogêneos

Ana Carolina Zoghbi

CEPESP – FGV e EESP – FGV

Paula Louzano

Departamento de Educação – USP e
Fundação Lemann

RESUMO

O objetivo deste trabalho é estimar o impacto do Caderno de Apoio e Aprendizagem – implementada em 2010 pela Secretaria Municipal de Educação de São Paulo – nas proficiências dos alunos da 4ª série do Ensino Fundamental. As evidências apontadas pela literatura sugerem que a aplicação de *Textbooks* em sala de aula impactam positivamente nos indicadores educacionais. Este estudo se diferencia dos demais por utilizar um painel de alunos com dados da Prova São Paulo de 2008 e de 2010, e também por estimar o impacto de um tratamento homogêneo, em que uma única intervenção (Caderno de Apoio) foi aplicada na mesma rede de ensino por um período de tempo semelhante (no mínimo três vezes na semana). Primeiramente calculamos o efeito médio do Caderno de Apoio nas notas por Primeiras Diferenças. Em seguida, estimamos por Regressão Quantílica o efeito heterogêneo do Caderno de Apoio ao longo da distribuição de conhecimento acumulado entre a 2ª e a 4ª séries. Os resultados sugerem que o Caderno Apoio tem efeito positivo sobre as proficiências, e esse efeito é heterogêneo ao longo da distribuição. Para Matemática, os impactos são positivos somente para os quantis superiores. Diferentemente, o Caderno de Apoio de Leitura apresenta um impacto considerável na nota nos quantis inferiores. Tal diferença pode ser explicada pelo fato do aprendizado em matemática estar mais associado aos conhecimentos adquiridos na escola, enquanto o aprendizado em leitura está também associado ao hábito de leitura, fortemente influenciado por características familiares. Por fim, foi realizado um teste de robustez (explorando uma falseabilidade) que sinaliza que não há auto-seleção dos alunos, o que confere maior confiabilidade aos resultados até agora encontrados.

Palavras-chave: Resultados Educacionais, Efeito do Tratamento, Primeiras Diferenças, Regressão Quantílica; Caderno de Apoio e Aprendizagem;

ABSTRACT

The aim of this paper was to estimate the impact of the textbooks *Caderno de Apoio e Aprendizagem* – implemented in 2010 by the Municipal Department of Education of São Paulo – on the test scores of students 4th graders. The evidence described in literature suggest that the use of textbooks in the classroom have a strong impact on educational indicators. This study innovates by using a panel of students with data from 2008 and 2010 editions of *Prova São Paulo*, and also by estimating the impact of a homogeneous treatment, i.e., a single intervention (*Caderno de Apoio*) that was applied to the same school system for a similar time period (at least three times a week). First, we calculate the average treatment effect of *Caderno de Apoio* on the test scores by First Differences. Next, we estimate the heterogeneous effect along the distribution of accumulated knowledge between the 2nd and 4th grades by Quantile Regression. The results show that the *Caderno de Apoio* has positive effects on the test scores, and this effect is heterogeneous over the distribution. For Math, the impact is positive only for the upper quantiles. In contrast, the *Caderno de Apoio* for reading has a considerable impact on test scores in lower quantiles. Such difference might be explained by the fact that math learning is closely tied to knowledge acquired at school, whereas reading learning is also associated to the reading habit, which is strongly influenced by family characteristics. Finally, we performed a robustness test (exploring falsifiability) that indicates there is no self-selection of students, which gives greater reliability to the results so far found.

Keywords: Educational Outcomes, Treatment Effect, First Differences, Quantile Regression, Textbooks,

Classificação JEL: I21; I28

Área ANPEC: Área 11 - Economia Social e Demografia Econômica

1. Introdução

Nos últimos anos o Brasil alcançou a universalização técnica do atendimento escolar entre pessoas com idade para cursar o Ensino Fundamental¹. Com esse aumento na margem extensiva da Educação, e com o surgimento de avaliações educacionais nacionais e internacionais², ganhou destaque o aumento da educação na margem intensiva, isto é, as questões relacionadas à melhoria de aprendizado dos alunos.

Dentre os fatores já identificados na literatura nacional e internacional que poderiam ampliar a educação na margem intensiva, merece destaque o papel do background familiar. De acordo com esses estudos, a influência da família na formação das habilidades de seus filhos é tão importante que o espaço para a atuação do setor público parece ser bastante limitado. Essa perspectiva é particularmente problemática para o Brasil, pois grande parte da população adulta tem baixa escolaridade, o que prejudica a transferência de conhecimento para formação de habilidades de seus filhos e perpetua o ciclo da pobreza.

Apesar do alcance limitado das políticas públicas para a melhoria de aprendizado dos alunos, alguns fatores não diretamente ligados à família se revelaram importantes de acordo com estudos empíricos, como por exemplo, a melhora da infra-estrutura escolar, qualificação de professores e diretores, *accountability*, etc.^{3,4,5} Dentre esses fatores, destaca-se o uso de “apostilas” (ou *textbooks*) em sala de aula pelos professores.

O objetivo deste trabalho é justamente estimar o impacto de uma nova apostila, denominada Caderno de Apoio e Aprendizagem (doravante denominado **Caderno de Apoio**), nas proficiências dos alunos da 4ª Série do Ensino Fundamental da rede Municipal de Educação de São Paulo. Apesar de já existirem estudos que buscam avaliar o impacto desse tipo de material no desempenho acadêmico⁶, este estudo se diferencia dos demais ao aproveitar uma estrutura de dados que permite acompanhar os alunos ao longo do tempo, e também por estimar o impacto de um tratamento homogêneo, em que uma única intervenção (Caderno de Apoio) foi aplicada a mesma rede de ensino por um período de tempo semelhante (no mínimo três vezes na semana). Para alcançar o objetivo deste trabalho, primeiramente calculou-se o efeito médio do Caderno de Apoio nas proficiências dos alunos pela metodologia de Primeiras Diferenças. Além disso, este estudo inova ao aplicar Regressão Quantílica para captar o efeito heterogêneo do Caderno de Apoio ao longo da distribuição de conhecimento acumulado entre a 2ª e a 4ª série.

Jamison, Serle e Galda (1981) encontram evidências de que o uso de *textbooks* em sala de aula pelos professores impacta positiva e significativamente nas proficiências dos alunos da Nicarágua. Comparativamente a outras alternativas de políticas públicas educacionais, o uso de apostilas apresenta claras vantagens. Por exemplo, um aumento na qualificação de professores ou uma redução do tamanho de turma apresentam efeitos inferiores aos efeitos estimados dos *textbooks*. Resultados semelhantes também foram encontrados por Heyneman, Farrel e Sepulveda-Stuardo (1978), mas os autores chamam atenção para o fato de que mesmo que o material didático

¹ A taxa de atendimento entre crianças de 7 a 14 anos era igual a 96,9% em 2010, o que pode ser considerado praticamente “universalização”, que seria igual a 98% de atendimento de acordo com o “Compromisso Todos pela Educação”. Vale ressaltar que 100% de atendimento é algo difícil de ser atingido devido a razões que fogem do alcance de políticas públicas (problemas de saúde, por exemplo).

² Pode-se citar como avaliações nacionais, o SAEB (Sistema Nacional de Avaliação da Educação Básica), a Prova Brasil, entre outras avaliações estaduais e municipais. Em relação às avaliações internacionais, o Brasil participa do PISA (*Programme for International Student Assessment*).

³ Sobre os resultados da *accountability* educacional ver Jacob (2005) e Carnoy e Loeb (2002).

⁴ Para saber mais sobre os efeitos das características dos professores no sucesso escolar ver Rivkin, Hanushek e Kain (2005).

⁵ Sobre os efeitos de algumas variáveis relacionadas a infra-estrutura escolar no desempenho dos alunos em exames padronizados ver Albernaz, Ferreira e Franco (2002).

⁶ Heyneman, Farrel e Sepulveda-Stuardo (1978); Fuller (1986); Jamison, Serle, Galda (1981); Fuller e Clarke (1994); Leme *et al.* (2012); entre outros.

seja um dos fatores de maior impacto no aprendizado, é necessário conhecer por quais mecanismos essa melhoria ocorre, uma vez que esse efeito não foi uniforme entre os países analisados.

Essa falta de uniformidade nos resultados entre países fica clara comparando estudos de diferentes países como em Glewwe *et ali.* (2007) – para o Quênia - e Leme *et ali.* (2012) – para o Brasil. Os autores primeiramente avaliam o efeito médio do *textbooks* na proficiência, e em seguida avaliam se esse impacto varia entre os alunos que apresentaram piores ou melhores desempenhos anteriormente a utilização da apostila. O primeiro trabalho não encontra impacto na média, mas evidencia que o material aumenta a proficiência daqueles que já apresentavam uma maior nota antes da implementação do *textbooks*. Já o segundo trabalho evidencia um impacto na média, e este efeito é maior entre os municípios que antes da implementação apresentavam um pior desempenho.

Alguns estudos também realizam uma análise de custo efetividade do *textbooks*. Nesse sentido, Lockheed e Hanushek (1987)⁷ encontram evidências de que o *textbooks* é um dos insumos com melhor custo efetividade. Os autores também chamam atenção para o fato de que no Brasil, esse resultado foi encontrado quando comparado com a qualificação de professores e com outros insumos de infraestrutura. Pritchett e Fimer (1997) afirmam que políticas públicas que direcionam recursos diretamente aos professores (como por exemplo, o aumento de salários), mesmo que o objetivo seja a melhoria da qualidade da educação, apresentam um custo efetividade pior comparativamente às políticas que direcionam recursos para insumos como livros, ou material de instrução. Tan, Lane e Lassibile (1999) encontram resultados que evidenciam um melhor custo efetividade para os *textbooks*, enquanto a análise realizada para merenda escolar resultou em um baixo custo efetividade. Os autores ressaltam que a análise de custos diretos e indiretos (custo de oportunidade) é importante para avaliar a efetividade de políticas públicas.

Os resultados deste trabalho evidenciam que o Caderno Apoio impacta positiva e significativamente nas proficiências de Leitura e de Matemática dos alunos da 4ª série do Ensino Fundamental da rede municipal de São Paulo. No entanto, esse impacto é heterogêneo ao longo da distribuição de conhecimento acumulado entre a 2ª e a 4ª série. Para a disciplina de matemática, o impacto do Caderno de Apoio na proficiência é positivo e significativo somente a partir do 5º quantil, e apresenta um maior impacto no 9º quantil. Diferentemente, o Caderno de Apoio de Leitura apresenta um impacto na nota considerável nos quantis inferiores, com destaque para o 1º quantil. Dessa forma, o Caderno de Apoio de Matemática apresenta maior efeito entre os alunos com maior acúmulo de conhecimento entre a 2ª e a 4ª série, isto é, entre aqueles com maior habilidade, e o Caderno de Apoio de Leitura apresenta maior impacto entre os alunos com menor habilidade (menor acúmulo de conhecimento entre a 2ª e a 4ª série). Por fim, foi realizado um teste de robustez (explorando uma falseabilidade) que sinaliza que não há auto-seleção dos alunos, o que confere maior confiabilidade aos resultados até agora encontrados.

Este trabalho está dividido em cinco seções, incluindo esta introdução. Na segunda seção é realizada uma apresentação dos dados da Prova São Paulo, no qual se encontram as informações do programa “Cadernos de Apoio e Aprendizagem” da rede Municipal de Educação de São Paulo, com algumas estatísticas descritivas. A seção 3 apresenta a justificativa para o uso de Primeiras Diferenças e Regressões Quantílicas como estratégias de estimação. A seção 4 apresenta os resultados das estimções e dos testes de robustez para verificar se a auto seleção do aluno foi realmente eliminada com as estratégias de estimação utilizadas. Finalmente na seção 5 são apresentadas as considerações finais.

2. Base de Dados

O banco de dados utilizado para este trabalho foi construído usando os microdados da Prova São Paulo. Esta prova é anual, e sua primeira aplicação ocorreu em 2007 nas escolas da rede

⁷ Os autores analisam quais políticas educacionais apresentam maior custo efetividade para um conjunto de países.

municipal da cidade de São Paulo. A aplicação é censitária para os alunos das séries pares (2^a, 4^a, 6^a e 8^a), e amostral para os alunos das séries ímpares (3^a, 5^a, 7^a)⁸. O principal objetivo da Prova São Paulo é avaliar o desempenho em Leitura e Matemática dos alunos, e com estes resultados subsidiar a secretaria de educação na formulação de políticas públicas para a melhoria da educação. Além de possuir informações relacionadas com as proficiências dos alunos, também abrange informações socioeconômicas dos próprios alunos, professores, diretores, supervisores e coordenadores pedagógicos. Adicionalmente, estes também respondem questões sobre o funcionamento, os problemas e a infraestrutura da escola.

No início do ano letivo de 2010, a Secretaria de Educação Municipal de São Paulo implementou um novo recurso didático, o Caderno de Apoio e Aprendizagem⁹. O Caderno de Apoio é elaborado tanto para a disciplina de Leitura, como para a de Matemática. Para obter mais informações sobre essa “apostila”, a Secretaria incluiu nos questionários da Prova São Paulo de 2010 dos alunos e dos professores questões sobre o uso do Caderno de Apoio.

Com essas informações adicionais coletadas em 2010 é possível avaliar como essa nova “apostila”, o Caderno de Apoio, afetou o aprendizado dos alunos pertencentes à Rede Municipal de Educação de São Paulo. Adicionalmente, o fato de podermos acompanhar os alunos ao longo do tempo possibilita a construção de um painel de alunos, em que no primeiro período nenhum aluno foi exposto ao Caderno de Apoio (2008), enquanto no segundo período somente uma parte dos alunos foi exposta ao Caderno de Apoio (2010).

A escolha do período, 2008 e 2010, deveu-se ao fato de que somente as séries pares são censitárias¹⁰. Por essa razão, é necessário pular um ano para acompanhar o máximo de alunos possíveis. Assim, as possibilidades de escolhas das séries a serem analisadas são: 2^a série (2008) e 4^a série (2010); ou 6^a série (2008) e 8^a série (2010). Optou-se pelo painel de alunos da 2^a série (2008) e da 4^a série (2010) para analisar o impacto da implementação do Caderno de Apoio nas proficiências de leitura e de matemática¹¹. Todas as estatísticas descritivas e exercício econométricos realizados são baseados neste painel.

Na Tabela 1 encontram-se as proficiências de matemática e de leitura para a 2^a série (2008) e a 4^a série (2010) para alguns pontos da distribuição. Observa-se que as notas médias são semelhantes às medianas. Além disso, as proficiências são maiores no percentis superiores. Em termos de conhecimento acumulado, o 99^o percentil apresenta um maior acúmulo, isto é, os alunos dos percentis superiores são aqueles com maior habilidade. Na tabela 2 encontram-se algumas características dos alunos de acordo com a posição na distribuição de proficiências de 2010. É possível observar nos percentis superiores de notas uma maior proporção de alunos com melhor *background* familiar.

⁸ A Prova São Paulo é baseada na escala SAEB, a qual utiliza a Teoria da Resposta ao Item (TRI).

⁹ O objetivo do caderno é auxiliar o docente durante suas aulas, de forma que eles sigam as orientações curriculares de aprendizado de cada série, e também cumpram o cronograma de ensino previsto. Além disso, permite uma maior otimização do tempo extraclasse dos professores gasto na elaboração das aulas, o que permite que os professores utilizem esse tempo para elaboração de atividades complementares. Esse material se diferencia do livro didático fornecido pelo Governo Federal pelo fato de “contemplar as especificidades da rede pública municipal paulistana do ponto de vista de suas realidades regionais, das condições de acervo de livros, de equipamentos disponíveis, de espaços físicos das escolas e do processo de formação de educadores desenvolvido nos últimos anos”. (<http://portalsme.prefeitura.sp.gov.br/default.aspx>).

¹⁰ A avaliação nas séries ímpares é amostral. Assim, 35 alunos de todas as séries ímpares são sorteados em cada escola para realizar a Prova São Paulo.

¹¹ Não seria possível analisar os alunos entre a 4^a série (2008) e a 6^a série (2010) porque o questionário respondido pelos alunos do Ciclo I (1^a a 4^a série) apresenta menos questões do que o respondido pelos alunos do ciclo II (5^a série a 9^a série). Em relação ao painel que acompanha os alunos da 6^a a 8^a série, espera-se que haja um viés de seleção, uma vez que as séries finais do Ensino Fundamental apresentam uma maior taxa de repetência. Dessa forma, os alunos que chegam a 8^a série são aqueles com maior habilidade.

Tabela 1 – Distribuição das Proficiências de Matemática e de Português

		Média	Desvio	25°	Mediana	75°	99°
Matemática	2ª série	142.26	33.49	117.81	139.27	164.51	227.80
	4ª série	181.03	44.49	149.55	181.22	212.31	288.40
Conhecimento Acumulado		38.77	-	31.74	41.94	47.80	60.59
Português	2ª série	134.01	37.15	106.68	134.33	160.06	218.59
	4ª série	170.00	44.68	140.26	169.52	200.21	275.19
Conhecimento Acumulado		36.00	-	33.58	35.19	40.15	56.66

Fonte: Elaboração Própria com os Microdados da Prova São Paulo de 2008 e de 2010.

Tabela 2 – Características dos Alunos segundo a Distribuição das Proficiências de Matemática

	Média (%)	Desvio (%)	25° Percentil (%)	Mediana (%)	75° Percentil (%)	99° Percentil (%)
Mãe Branca	40.01	0.49	37.35	41.94	46.94	51.07
Pai Branco	38.07	0.49	35.86	39.76	43.91	47.50
Mãe com Ensino Médio	21.50	0.41	17.49	24.63	30.43	35.53
Pai com Ensino Médio	17.18	0.38	13.78	19.80	24.34	29.15
Mãe desempregada	25.11	0.43	26.68	23.97	22.46	21.51
Pai desempregado	8.75	0.28	9.66	8.13	7.11	5.97
Renda familiar entre R\$ 2126 e R\$ 4250	3.23	0.18	2.72	3.58	4.94	8.06
Renda familiar maior que R\$ 4250	0.79	0.09	0.71	0.84	0.95	1.25

Fonte: Elaboração Própria com os Microdados da Prova São Paulo de 2010.

No que diz respeito ao uso do Caderno de Apoio, existem algumas perguntas sobre a intensidade de uso nos questionários dos professores e dos alunos. Primeiramente é perguntado se o professor usa ou não o Caderno de Apoio. Em seguida, qual é a intensidade de uso: 1 a 2 vezes por semana; 3 a 4 vezes por semana; todos os dias da semana. A recomendação da Secretaria Municipal de Educação de São Paulo foi a de que os professores usassem o material mais do que três vezes na semana. Dessa forma, o Caderno de Apoio auxiliaria os professores a seguirem as orientações curriculares de cada série, e também a cumprirem o cronograma de ensino previsto. Jamison *et al.* (1981) argumentam que muitas intervenções educacionais consistem na provisão de materiais didáticos para uso em sala de aula, mas seus impactos na melhoria da educação dependerão da forma como são explorados.

Por essa razão, optamos por construir uma *dummy* de intensidade de uso do material. Assim, os alunos que foram expostos ao Caderno de Apoio no mínimo 3 vezes por semana apresentam uma *dummy* igual 1, e caso sejam exposto menos do que 3 vezes por semana, ou não tiverem nenhuma exposição ao Caderno de Apoio, a *dummy* será igual a 0. Vale destacar que isso permite uma homogeneização do tratamento (exposição ao Caderno de Apoio), uma vez que todos os alunos que foram expostos ao Caderno de Apoio no mínimo três vezes na semana apresentam maior probabilidade de terem sido submetidos ao mesmo conteúdo. Isso reforça a idéia de que o Caderno de Apoio é um tratamento homogêneo. Pela Tabela 3, é possível notar que há uma concentração das respostas dos professores e os alunos nas alternativas “1 a 2 vezes por semana” e “3 a 4 vezes por semana”. Adicionalmente, observa-se que a proporção de professores e de alunos que reportaram usar intensamente o Caderno de Apoio, como principal material, é maior do que aqueles que reportaram não usar intensamente o Caderno de Apoio (material secundário).

Tabela 3 – Proporção de Professores e de Alunos que reportaram o uso dos Cadernos de Apoio (%)

	Português		Matemática	
	Professores (%)	Alunos (%)	Professores (%)	Alunos (%)
Não usa	1.02	5.3	1.19	5.62
Usa de 1 a 2 vezes	41.61	32.55	40.89	33.62
Usa de 3 a 4 vezes	49.68	39.72	50.55	39.18
Usa todos os dias	7.61	22.43	7.35	21.58
Usa Sempre Caderno (Principal)	57.29	62.15	57.90	60.76
Não Usa Sempre Caderno (Secundário)	42.63	37.85	42.08	39.24

Fonte: Elaboração Própria com os Microdados da Prova São Paulo de 2010.

Em relação às características dos professores que optaram ou não pelo uso do Caderno de Apoio como principal material didático dentro de sala de aula (Tabela 4), observa-se uma diferença estatisticamente significativa para experiência do professor e distância da escola. Pelo fato de existir essas diferenças entre os professores que usaram e aqueles que não usaram intensamente o Caderno de Apoio, é necessário controlar pelas características dos professores nas estimações da seção 4.

Tabela 4 – Características dos Professores segundo a intensidade do uso do Caderno de Apoio (%)

	Usa muito o Caderno Apoio de Matemática (%)		Teste de Igualdade de Médias	Usa muito o Caderno Apoio de Matemática (%)		Teste de Igualdade de Médias
	Não	Sim		Não	Sim	
prof_superior	0.776	0.759	Aceita	0.772	0.761	Aceita
prof_especializacao	0.212	0.235	Aceita	0.216	0.233	Aceita
prof_mestrado	0.012	0.005	Aceita	0.012	0.005	Aceita
prof_doutorado	0.000	0.001	Aceita	0.000	0.001	Aceita
prof_tempomenos1	0.010	0.006	Aceita	0.011	0.005	Aceita
prof_tempo1a5anos	0.061	0.060	Aceita	0.064	0.058	Aceita
prof_tempo6a10anos	0.121	0.109	Aceita	0.123	0.107	Aceita
prof_tempo11a15anos	0.202	0.153	Rejeita***	0.202	0.152	Rejeita***
prof_tempo16a20anos	0.229	0.210	Aceita	0.226	0.211	Aceita
prof_tempomais20anos	0.377	0.463	Rejeita***	0.373	0.467	Rejeita***
prof_moraperto	0.679	0.719	Rejeita**	0.684	0.716	Aceita

Fonte: Elaboração Própria com os Microdados da Prova São Paulo de 2010. ***1%; **5%, *10%.

Por fim, no quadro 1 encontram-se as variáveis que foram usadas nas estimações. Ao inserir características dos alunos e dos professores, é possível corrigir o viés de seleção gerado por observáveis que variam ao longo do tempo. Para captar um possível efeito de *peer effect*, foram construídas duas variáveis: média de colegas cujos pais ajudam bastante a fazer a lição; média de colegas que entraram na 1ª série. Da mesma forma, seria interessante captar como a redução no tempo de aprendizado impacta na proficiência do aluno. Por isso, duas variáveis (o aluno falta muito a aula; o professor perde mais da metade do tempo para começar a aula) foram incorporadas no modelo. Além disso, algumas variáveis de percepção do aluno (aprendo bastante) e do professor (professor acredita que mais de 80% da turma será aprovada; professor considera ter bastante violência na escola) serviram de controle. Assim como algumas características dos alunos que pudessem variar ao longo do tempo (tem muita lição; professor corrige lição; pais ajudam bastante a fazer a lição). Pelo fato do professor do aluno não ser o mesmo nos dois anos do painel, informações sobre escolaridade, experiência e distância da casa até a escola também fizeram parte dos modelos estimados.

Quadro 1 – Descrição da Variável Utilizadas nas Estimções Econométricas

Variáveis	Descrição
<i>usa sempre CA</i>	<i>Dummy igual a 1 se o professor usa material mais do que 3 dias na semana, 0 caso contrário</i>
tem muita lição	<i>Dummy igual a 1 se o aluno faz muita lição, e 0 caso contrário</i>
pais ajudam bastante a fazer a lição	Dummy igual a 1 se os pais sempre ajudam o aluno a fazer a lição, e 0 caso contrário
Falta aula	Dummy igual a 1 se o aluno falta muito a aula, e 0 caso contrário
professor corrige lição	Dummy igual a 1 se o professor do aluno corrige a lição sempre, e 0 caso contrário
aprendo bastante	<i>Dummy igual a 1 se o aluno tem a percepção que aprende bastante, e 0 caso contrário</i>
média de colegas cujos pais ajudam bastante a fazer a lição	Proporção de colegas do aluno cujos pais ajudam na lição
média de colegas que entraram na 1ª série	Proporção de colegas do aluno que entraram na escola na 1ª série do Ensino Fundamental
professor frequentou mais que ES	Dummy igual a 1 se o professor frequentou mais que o Ensino Superior, e 0 caso contrário
professor tem mais de 10 anos de experiência	Dummy igual a 1 se o professor tem mais de 10 anos de experiência e 0 caso contrário
perco mais da metade da aula para começar a aula	Dummy igual a 1 se o professor perde mais da metade da aula para começar a aula, e 0 caso contrário
professor mora perto	Dummy igual a 1 se o professor mora perto da escola, e 0 caso contrário
professor acredita que mais de 80% da turma será aprovada	Dummy igual a 1 se o professor respondeu que mais de 80% da turma será aprovada, e 0 caso contrário
professor considera ter bastante violência na escola	Dummy igual a 1 se o professor considera a escola bastante violenta, e 0 caso contrário

Fonte: Elaboração Própria.

3. Estratégia de Estimação

Em linhas gerais, este trabalho se baseia em duas hipóteses. A primeira é a de que o uso do Caderno de Apoio impacta positivamente nas proficiências dos alunos, mas esse impacto é heterogêneo ao longo da distribuição de conhecimento acumulado entre a 2ª e a 4ª série. Para testar estas hipóteses, estimamos por Primeiras Diferenças o efeito médio do tratamento (exposição ao Caderno de Apoio) na proficiência dos alunos da 4ª série em 2010. Em seguida, estimamos o impacto do tratamento ao longo da distribuição de conhecimento acumulado entre a 2ª e a 4ª séries por meio de Regressão Quantílica.

1.1. Estimação por Primeiras Diferenças (FD)

Como mencionado anteriormente, o objetivo deste trabalho é verificar o efeito do Caderno de Apoio (tratamento) na proficiência dos alunos da 4ª série das escolas da rede municipal da cidade de São Paulo em 2010. No entanto, os alunos expostos ao Caderno de Apoio podem apresentar características não observáveis que os levam a alcançar melhores resultados nas provas. Por essa razão, é necessário escolher um método de estimação que corrija esse possível viés de seleção. Um dos métodos quasi-experimentais mais eficazes para corrigir esse viés é o de Primeiras Diferenças (FD). Esse método corrige o problema de variáveis omitidas fixas no tempo correlacionadas com a variável de tratamento, e também a auto-seleção por observáveis que variam no tempo.

O efeito médio do tratamento é dado pela seguinte expressão:

$$E[Y_i^{CA} - Y_i^{NCA}] \quad (1)$$

Em que Y_i^{CA} é a proficiência esperada para o aluno i que foi exposto ao Caderno de Apoio, e Y_i^{NCA} é a proficiência esperada para o mesmo aluno i caso ele não tivesse sido exposto ao Caderno de Apoio (NCA).

No entanto, não é possível observar Y_i^{CA} e Y_i^{NCA} no mesmo período. Só observamos os resultados de quem foi ou não foi tratado:

$$E[Y_i^{CA}|CA] - E[Y_i^{NCA}|NCA] \quad (2)$$

Ao somar e subtrair o valor esperado contrafactual dos não tratados $E[Y_i^{NCA}|CA]$ em (2), obtemos a seguinte expressão:

$$E[Y_i^{CA} - Y_i^{NCA}|CA] + \{E[Y_i^{NCA}|CA] - E[Y_i^{NCA}|NCA]\} \quad (3)$$

Em que primeiro termo é o efeito do tratamento sobre os tratados (ATT), e o segundo termo refere-se ao viés de seleção.

Segundo Rubin (1977), ao condicionar nas características observáveis (X) que determinam a seleção ao tratamento CA, os resultados potenciais independem do *status* do tratamento. Dessa forma, temos que $E[Y_i^{NCA}|CA, X] = E[Y_i^{NCA}|NCA, X]$ e o ATT pode ser estimado consistentemente se condicionado em características observáveis, produzindo o ATT(x). A escolha das variáveis observáveis que determinam a seleção do tratamento foi baseada nas características dos alunos e dos professores que variam ao longo do tempo.¹²

Uma vez equacionado o problema de auto-seleção por observáveis que variam no tempo, aplicamos o método de Primeiras Diferenças. Por esse método, além de obtermos estimadores robustos à auto-seleção por observáveis, i.e., características que determinam a auto-seleção de alunos e professores que ao mesmo tempo influenciam o desempenho dos estudantes e a propensão a adotar o Caderno de Apoio por parte dos docentes; também corrigimos potenciais vieses associados a características não observáveis fixas no tempo e correlacionadas com a variável de tratamento. Desse modo, é um método menos restritivo e, portanto, mais desejável.

O modelo a ser estimado pelo método de Primeiras Diferenças (FD) é representado pela seguinte equação:

$$Y_{it} = \beta_0 + \beta_1 CA_i + \delta CA_i T_t + \beta_2 T_t + X_{it} \beta_3 + c_i + \varepsilon_{it} \quad (4)$$

em que Y_{it} representa as proficiências do aluno i no período t (2008 e 2010), o termo X é um vetor linha de covariadas com as características dos alunos e dos professores, e β_3 representa seu respectivo vetor coluna de parâmetros. O termo c_i representa as características não observáveis fixas no tempo que podem ou não estar correlacionadas com os demais regressores; T é uma *dummy* de ano igual a 0 em 2008, e igual a 1 em 2010; e ε_{it} é o termo aleatório de erro não correlacionado com os regressores ou com o efeito fixo. As características dos alunos e dos professores incluídas no modelo encontram-se no Quadro 1 da seção anterior.

A interação $CA_i T_t$ refere-se à variável de tratamento e assume valor igual a 0 em 2008 para todos os alunos. Como o tratamento referente ao uso Caderno de Apoio ocorreu no início de 2010, nesse ano, observa-se que o termo cruzado $CA_i T_t$ é igual a 1 somente para alguns alunos que sofreram a intervenção. Considerando a estrutura dos dados, a transformação realizada em (4) ao se tomarem as primeiras diferenças (FD) resulta em uma cross-section de dados em que o efeito não

¹² Essas variáveis foram definidas na seção anterior.

observado fixo no tempo é eliminado, e em que a variável de tratamento degenera em CA_{it} , igual a 1 para os alunos que foram expostos ao Caderno de Apoio e 0 caso contrário. Portanto, temos a equação (4) transformada:

$$\Delta Y_{it} = \beta_3 + \delta CA_{it} + \Delta X_{it} \beta_2 + \Delta \varepsilon_{it} \quad (5)$$

em que ΔY_{it} representa a diferença entre as proficiências do aluno i nos períodos t e $t - 1$, isto é, o aprendizado dos alunos entre 2008 (2ª série) e 2010 (4ª série).

Pela equação 5, observamos que somente as variáveis que variam no tempo mantêm-se no modelo. Após a transformação do modelo por Primeiras Diferenças, podemos aplicar estimar o modelo por Mínimos Quadrados Ordinários¹³ para obtermos o impacto médio do Caderno de Apoio, δ .

1.2. Estimação do Efeito do Tratamento na Distribuição dos Resultados Potenciais

Angrist e Pischke (2009) argumentam que quando a variável dependente apresenta uma distribuição contínua, esta pode mudar de várias maneiras que não poderiam ser reveladas por uma estimação na média. Como a variável dependente deste trabalho é contínua, e existem algumas evidências de que as características observáveis mudam de acordo com a posição na distribuição de proficiências dos alunos (Tabelas 1 e 2 da seção anterior), estimamos também o impacto do tratamento por meio da Regressão Quantílica. Nossa segunda hipótese é a de que o efeito do tratamento (exposto sempre ao Caderno de Apoio) pode ser heterogêneo ao longo da distribuição.

Supondo que o modelo seja linear, e aplicando o modelo de Regressão Quantílica nas variáveis em primeira diferença, de modo semelhante à equação 5, temos que:

$$\Delta Y_{it}^{CA} = \beta_3^\tau + \delta^\tau CA_{it} + \Delta X_{it} \beta_2^\tau + \varepsilon_{it} \quad \text{e} \quad Q_{\varepsilon_{it}}^\tau = 0 \quad (6)$$

O subscrito i refere-se ao aluno, e $CA \in \{0,1\}$. $Q_{\varepsilon_{it}}^\tau$ representa o quantil $\tau - th$ da variável aleatória não observada, ε_i . β_3^τ , β_2^τ e δ^τ são os parâmetros não conhecidos desse novo modelo. δ^τ representa o *Quantile Treatment Effect* (QTE) condicional nas variáveis observáveis

Dessa forma, estamos avaliando o efeito do Caderno de Apoio ao longo da distribuição de aprendizado acumulado pelos estudantes entre a 2ª (2008) e 4ª (2010) séries do Ensino Fundamental. Note que o modelo de regressão quantílica aplicado às variáveis em primeiras diferenças não se origina de um modelo em níveis, tal como na equação (5) – estimada pelo método de Primeiras Diferenças. Isso decorre do fato de que ao se avaliar o desempenho em níveis, supõe-se um coeficiente constante ao longo do tempo. Mas no caso de um modelo com efeitos heterogêneos ao longo da distribuição, um indivíduo em um quantil de desempenho em um ano pode mudar de quantil no ano seguinte, e seus coeficientes seriam diferentes. Por isso deve-se interpretar os quantis como quantis de aprendizados acumulado entre os anos.

Para identificação do QTE, além da suposição de linearidade, é necessário supor que CA , ΔX sejam exógenos.

$$\varepsilon \perp (\Delta X, CA) \quad (7)$$

$$Q_{\Delta Y_{it} | \Delta X, CA}^\tau = \beta_3^\tau + \delta^\tau CA + \Delta X \beta_2^\tau \quad (8)$$

¹³ Se o período de análise fosse superior a dois, o estimador a ser aplicado seria o de Mínimos Quadrados Ordinários Ponderados. Para maiores detalhes ver Wooldridge (2010).

Ao estimar a equação (6) é possível recuperar os parâmetros desconhecidos dos resultados potenciais da distribuição conjunta de ΔY , ΔX e CA . Esses coeficientes desconhecidos podem ser estimados por Regressão Quantílica¹⁴.

$$(\hat{\beta}^\tau, \hat{\delta}^\tau) = \underset{\beta_2, \beta_3, \delta}{\operatorname{argmin}} \sum \rho_\tau(\Delta Y_{it} - \Delta X_i \beta_2 - CA_i \delta - \beta_3) \quad (9)$$

$$\rho_\tau(u) = u \cdot \{\tau - 1(u < 0)\} \quad (10)$$

Com parâmetros estimados recuperamos a distribuição dos tratados $F_{CA|CA==1}^{-1}$ e dos controles $F_{NCA|CA==1}^{-1}$, o que nos permite calcular o efeito de tratamento para cada quantil (QTE):

$$QTE(\tau) = F_{CA}^{-1}(\tau) - F_{NCA}^{-1}(\tau)$$

4. Análise dos Resultados

Os resultados das estimações por Efeito Fixo encontram-se nas Tabelas 5 e 6 para as amostras de Matemática e de Leitura, e servirão de *baseline* para as estimações por Regressão Quantílica nas Tabelas 7 e 8. O modelo FE1 representa a estimação do efeito médio do Caderno de Apoio de matemática e de leitura nas proficiências dos alunos da 4ª série sem variáveis de controle. Os modelos FE2 e FE3 são as estimativas por Efeito Fixo do impacto do Caderno de Apoio ao inserir os controles dos alunos e dos professores.

Os resultados das Tabelas 5 e 6 sugerem que o Caderno de Apoio, tanto de Matemática, como de Leitura, apresenta impacto médio positivo nas proficiências. Esse efeito reduz ligeiramente quando incluímos as variáveis de controle, apresentado menor impacto quando inseridos os controles dos alunos e dos professores conjuntamente (FE3). Os impactos do Caderno de Apoio no modelo FE3 são de 1.89 pontos em Matemática e de 2.44 pontos em leitura. Isso representa, aproximadamente, 6% e 7% dos desvios padrão das proficiências médias de matemática e de leitura iniciais dos alunos¹⁵. Esses resultados corroboram os resultados encontrados por Leme, *et al.* (2010), Lockheed e Hanushek (1987) e Jamison *et al.* (1981), que avaliam o impacto de *textbooks* na melhoria do desempenho dos alunos em países em desenvolvimento.

Como mencionado na seção anterior, um dos pontos que diferencia este trabalho dos demais trabalhos é que este avalia o impacto de “apostilas” nos indicadores educacionais usando um painel de alunos. Por isso, seria interessante revisitar os resultados de algumas variáveis de controle usualmente utilizadas na literatura em outras estruturas de banco de dados (*cross section*; painel de municípios; painel de escolas). Em relação às características escolares e familiares dos alunos, é possível verificar que o fato do aluno ter mais lição de casa, ter pais que ajudem na lição, faltar aula, ter boa percepção de quanto aprende e ter amigos cujos pais ajudam com a lição, apresentaram o sinal esperado e foram estatisticamente significativas dependendo do modelo analisado.

O fato do aluno ter mais lição tem um efeito positivo, e sua magnitude aumenta ao controlar pelas características dos professores, tanto para as estimações de matemática (Tabela 5), como de leitura (Tabela 6).

No que se refere aos pais ajudarem na lição, na amostra de matemática, os resultados sugerem que quando são inseridos os controles dos professores, seu efeito diminui e perde significância. Já para as estimações em leitura, seu efeito aumenta.

Como se pode observar nas Tabelas 5 e 6, a variável falta aula apresenta um coeficiente de -3.74 e de -2.70 nas proficiências de matemática e de leitura, respectivamente. Da mesma forma, a variável que capta a percepção do aluno em relação ao que ele aprende apresenta um efeito negativo

¹⁴ Koenker e Bassett (1978) e Frölich e Melly (2010).

¹⁵ Os desvios padrão da proficiência de Matemática e de Leitura dos alunos da 2ª série em 2008 encontram-se na Tabela 1.

nas proficiências de matemática e de leitura, e perde magnitude no modelo FE3 quando comparado com o modelo FE2, que apresenta somente os controles dos alunos.

Em relação à proporção de colegas cujos pais ajudam bastante na lição, vale destacar que o intuito desta é captar em alguma medida o *peer effect* dos alunos. O impacto do coeficiente foi considerável, tanto para matemática (coeficiente de 9.17), como para leitura (coeficiente de 5.06). No entanto, são poucas as políticas públicas que capazes de influenciar as características das famílias.

Tabela 5 – Impacto médio na Proficiência de Matemática da 4ª série – Efeito Fixo

	FE1	FE2	FE3
usa sempre CA de Matemática	2.5538*** (0.4715)	2.2787*** (0.5888)	1.8959*** (0.6780)
tem muita lição		1.6445*** (0.4235)	1.6754*** (0.4892)
pais ajudam bastante a fazer a lição		1.1252** (0.4552)	0.9222* (0.5270)
falta aula		-3.7226*** (0.5266)	-3.7395*** (0.6084)
professor corrige lição		0.6210 (0.6210)	0.9646 (0.7206)
aprendo bastante		2.7238*** (0.7274)	1.8816** (0.8420)
média de colegas cujos pais ajudam bastante a fazer a		8.7756*** (1.5001)	9.1762*** (1.7298)
média de colegas que entraram na 1ª série		-1.8704 (1.8792)	-3.1752 (2.1741)
professor frequentou mais que ES			0.7922 (0.6891)
professor tem mais de 10 anos de experiência			0.9540 (0.7333)
perco mais da metade da aula para começar a aula			-2.2005*** (0.7264)
professor mora perto			-0.4591 (0.5523)
professor acredita que mais de 80% da turma será			3.9228*** (1.1682)
professor considera ter bastante violência na escola			-1.4366 (0.9590)
Constant	142.9545*** (0.1482)	137.6144*** (1.2103)	136.2586*** (2.0033)
Dummy	sim	sim	sim
R-squared	0.07	-0.04	-0.19
N	54997	42290	36762

Obs: Erros-padrão robustos entre parênteses. * p<0.10, ** p<0.05, *** p<0.01

No que diz respeito às características dos professores, estas apresentam o efeito esperado quando estatisticamente significativas. Na Tabela 5, os resultados mostram que os alunos cujos professores frequentaram mais do que o Ensino Superior, a nota em matemática é 1.15 pontos superior àqueles alunos cujos professores frequentaram no máximo o Ensino Superior.

Segundo o relatório do World Bank (2010), nos países da OCDE, 85% do tempo do professor em sala de aula é utilizado para atividades acadêmicas. Já no Brasil, somente 65% do tempo do professor em sala de aula é utilizado para atividades acadêmicas. Dado a existência desse desperdício do tempo do professor em sala de aula em atividades não acadêmicas, optou-se por inserir uma variável que avaliasse o impacto desse desperdício (perco mais da metade do tempo de aula para começar a aula) na nota. O resultado encontrado foi quase idêntico para matemática e leitura, uma redução na nota em 2.22 e em 2.20, respectivamente. Se somados a esses coeficientes

os impactos negativos dos alunos que faltam a aula, iguais a 3.74 e de 2.70 para matemática e leitura, respectivamente, a perda em termos de proficiência devido a fatores que reduzem o tempo gasto aprendendo pelos alunos seria nessa mesma ordem de aproximadamente 6.00 e 5.00 pontos.

Tabela 6 – Impacto médio na Proficiência de Leitura da 4ª série – Efeito Fixo

	FE1	FE2	FE3
usa sempre CA de Português	2.5859*** (0.4670)	2.6059*** (0.5954)	2.4404*** (0.6815)
tem muita lição		1.3310*** (0.4282)	1.3544*** (0.4905)
pais ajudam bastante a fazer a lição		0.9461** (0.4603)	1.0858** (0.5285)
Falta aula		-2.4612*** (0.5359)	-2.7008*** (0.6154)
professor corrige lição		0.0351 (0.6298)	-0.0160 (0.7248)
aprendo bastante		3.8302*** (0.7386)	3.6011*** (0.8459)
média de colegas cujos pais ajudam bastante a fazer a lição		4.1462*** (1.5147)	5.0600*** (1.7336)
média de colegas que entraram na 1ª série		-0.3383 (1.9004)	-2.4709 (2.1800)
professor frequentou mais que ES			1.1501* (0.6922)
professor tem mais de 10 anos de experiência			0.3766 (0.7376)
perco mais da metade da aula para começar a aula			-2.2204*** (0.7303)
professor mora perto			-0.6563 (0.5567)
professor acredita que mais de 80% da turma será aprovada			2.1747* (1.1727)
professor considera ter bastante violência na escola			-0.5822 (0.9647)
Constant	134.9592*** (0.1075)	131.9722*** (1.2233)	132.2014*** (2.0096)
Dummy de ano	sim	sim	sim
R-squared	0.61	-0.18	-0.35
N	53810	41102	35759

Obs: Erros-padrão robustos entre parênteses. * p<0.10, ** p<0.05, *** p<0.01.

Por fim, no que diz respeito às variáveis de controle, se a percepção do professor em relação à turma é positiva (acredita que mais de 80% da turma será aprovada), isso resulta em um impacto de 3.92 pontos em matemática e de 2.17 pontos em leitura, mas somente estatisticamente significativo a 10% em leitura.

A segunda hipótese testada neste trabalho é a de que o efeito do tratamento é heterógeno ao longo da distribuição. Para verificar essa hipótese, foi aplicada a regressão quantílica no modelo FE3, o qual já se encontra em Primeiras Diferenças. Dessa forma, é possível avaliar o impacto do Caderno de Apoio nos quantis da distribuição do conhecimento acumulado entre a 2ª e a 4ª série.

Os resultados da Tabela 7 mostram que o uso do Caderno de Apoio começa a ter impacto positivo e significativo na proficiência de matemática a partir do 5º quantil. Esse impacto varia de 2.13 (quinto quantil) a 2.92 (último quantil), e vai ao encontro da hipótese de que o caderno de apoio apresenta um efeito heterogêneo ao longo da distribuição do conhecimento acumulado entre a 2ª e 4ª série. Na Figura 1 é possível observar com mais detalhe os efeitos. Um maior impacto dos *textbooks* entre os alunos que já apresentavam maior nível de aprendizado também é encontrado por Glewwe *et al.* (2007).

Considerando os resultados do 9º e do 1º quantis, e seus respectivos erros padrão, notamos que o efeito no 9º quantil pode alcançar um impacto de aproximadamente 5 pontos¹⁶, comparativamente a um impacto nulo no 1º quantil.

Ao analisar o impacto do Caderno de Apoio de leitura na nota para cada quantil (Tabela 8), observamos que esse impacto é maior nos quantis inferiores, variando de 3.94, no 1º quantil, a um impacto nulo no último quantil. Nos cinco primeiros quantis, em que o conhecimento acumulado entre a 2ª e a 4ª série é menor, isto é, os alunos apresentam uma menor habilidade, o efeito do caderno é estatisticamente significativo a 1%, enquanto a partir do 6º quantil, o coeficiente perde significância. Esses resultados também podem ser observados na Figura 2.

Diferentemente dos resultados encontrados para o Caderno de Apoio de matemática (Tabela 7), o Caderno de Apoio de leitura (Tabela 8) apresenta um impacto na nota considerável nos quantis inferiores, com destaque para o 1º quantil cujo impacto é de 3.904. Considerando um limite superior de dois erros-padrão, esse resultado pode alcançar 6.6. Tal diferença pode ser explicada pelo fato do aprendizado em matemática estar mais associado aos conhecimentos adquiridos na escola, enquanto o aprendizado em leitura está também associado ao hábito de leitura, fortemente influenciado por características familiares.

As estimativas completas das Tabelas 7 e 8 podem ser verificadas nas tabelas A1 e A2 do Apêndice. Vale destacar algumas variáveis de controle que mantiveram o sinal e a significância encontrados nas estimações por Efeito Fixo (Tabelas 5 e 6) em quase todos os quantis. Analisando as estimações por regressão Quantílica para a proficiência de Matemática, o aluno que falta aula apresenta uma redução na nota, e esta varia de 5 pontos (1º quantil) a 3.5 pontos (9º quantil). Já a variável que capta em alguma medida o *peer effect* do aluno, média de colegas cujos pais ajudam bastante a fazer lição de casa, impacta positivamente na nota de matemática, principalmente nos quantis inferiores. Por fim, o professor que perde mais da metade do tempo em sala para começar a aula resulta em menores proficiências em todos os quantis avaliados. Vale ressaltar que em alguns quantis a redução da nota causada pelo desperdício de tempo é quase idêntica ao ganho gerado pelo Caderno de Apoio de Matemática.

Avaliando os resultados para Leitura na Tabela A2, é possível observar que o fato do aluno faltar aula resulta em uma diminuição na nota quase que na mesma magnitude que o Caderno de Apoio de Leitura melhora a nota para alguns quantis. Em relação à variável de *peer effect*, esta apresenta um impacto positivo e significativo em quase todos os quantis. Da mesma forma que observamos para a amostra de matemática. O professor que perde mais da metade do tempo em sala para começar a aula, reduz a nota do aluno em magnitudes que variam de 2.76 (1º quantil) a 1.61 (7º quantil).

Além das estimações com os controles de alunos e de professores, também foi estimada uma especificação adicional para as proficiências de matemática e de leitura cujos resultados encontram-se nas Tabelas 9, 10 e 11. Essa especificação adicional apresenta uma variável que busca identificar se aplicação de Primeiras Diferenças e os controles por observáveis foram suficientes para corrigir o problema de auto-seleção dos alunos. Essa variável, denominada de “placebo”, assume valor igual a 1 em 2010 para os alunos que foram expostos somente a um dos Cadernos de Apoio, Matemática ou Leitura, 0 para os alunos que foram expostos aos dois Cadernos simultaneamente, ou não foram expostos a nenhum dos Cadernos. A variável reflete, portanto, uma situação contrafactual. O intuito é verificar se os alunos que foram expostos somente a um dos Cadernos apresentam melhores notas na disciplina a qual não foram expostos ao Caderno. Dessa forma, se o aluno foi exposto somente ao Caderno de Apoio de Matemática, e também apresenta boas notas em leitura, seria um indicativo de auto seleção do aluno, isto é, os pais que fornecem melhor *background* familiar estarem colocando seus filhos em escolas com melhores práticas direcionadas para o ensino. Portanto, os coeficientes significativos da variável placebo indicariam a existência de auto seleção.

As estimativas completas dos modelos expostos nas Tabelas 9, 10 e 11 podem ser obtidas junto aos autores. A inclusão da variável placebo no modelo FE3 para a estimação por Efeito Fixo

¹⁶ O valor máximo que o coeficiente poderia obter ao somar o coeficiente com duas vezes o erro padrão do coeficiente.

Tabela 7 – Impacto do Caderno de Apoio de Matemática nos Quantis da Distribuição do Conhecimento Acumulado entre 2ª e 4ª série – Regressão Quantílica

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Intercepto	7.565*** (1.825)	20.826*** (1.333)	30.411*** (1.146)	39.371*** (1.094)	46.419*** (0.878)	54.650*** (1.104)	63.256*** (1.044)	73.777*** (1.142)	86.257*** (1.473)
Usa sempre Caderno de Apoio de Matemática	1.977 (1.331)	0.693 (0.987)	1.282 (0.931)	1.229 (0.850)	2.138*** (0.713)	2.474*** (0.709)	1.622** (0.7950)	1.432 (0.884)	2.923*** (1.054)
Controles: alunos e professores	sim	sim	sim	sim	sim	sim	sim	sim	sim

Obs: Erros-padrão robustos entre parênteses. * p<0.10, ** p<0.05, *** p<0.01

Tabela 8 – Impacto do Caderno de Apoio de Leitura nos Quantis da Distribuição do Conhecimento Acumulado entre 2ª e 4ª série – Regressão Quantílica

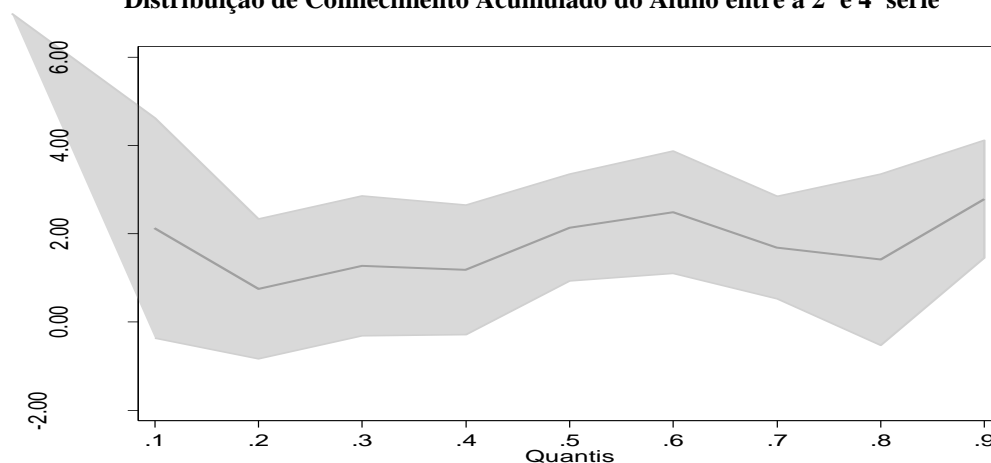
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Intercepto	-0.995 (1.737)	16.165*** (1.154)	25.406*** (1.043)	33.726*** (1.070)	41.444*** (1.083)	49.647*** (1.337)	58.553*** (1.296)	67.577*** (1.265)	81.719*** (1.895)
Usa sempre Caderno de Apoio de Leitura	3.904*** (1.346)	3.301*** (0.986)	3.260*** (0.938)	2.897*** (0.752)	2.023*** (0.782)	1.893** (0.854)	1.573* (0.805)	2.024** (0.898)	1.804 (1.323)
Controles: alunos e professores	sim	sim	sim	sim	sim	sim	sim	sim	sim

Obs: Erros-padrão robustos entre parênteses. * p<0.10, ** p<0.05, *** p<0.01

As estimativas completas dos modelos expostos nas Tabelas 9, 10 e 11 podem ser obtidas junto aos autores. A inclusão da variável placebo no modelo FE3 para a estimação por Efeito Fixo (Tabela 9), indica que não há auto seleção do aluno, já que os coeficientes da variável placebo para as amostras de matemática e de leitura não foram significativos.

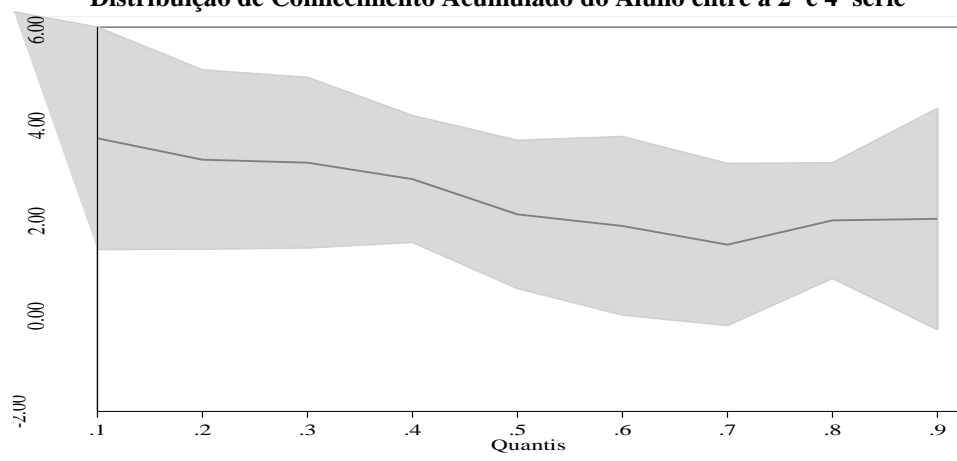
Nas estimações por Regressão Quantílica (Tabelas 10 e 11), as variáveis “placebo” não foram significativas em nenhum dos quantis dos modelos que avaliam os impactos do Caderno de Apoio em Matemática e Leitura, sugerindo inexistência de auto-seleção por parte dos alunos. Dessa forma, não há indícios de que o efeito do tratamento possa se dever ao fato de que pais de alunos com melhor *background* matriculam seus filhos em escolas que adotam as melhores práticas de ensino.

Figura 1 – Impacto do Caderno de Apoio de Matemática na Proficiência de Matemática ao Longo da Distribuição de Conhecimento Acumulado do Aluno entre a 2ª e 4ª série



Fonte: Elaboração Própria com os Microdados da Prova São Paulo de 2008 e 2010.

Figura 2 – Impacto do Caderno de Apoio de Português na Proficiência de Matemática ao Longo da Distribuição de Conhecimento Acumulado do Aluno entre a 2ª e 4ª série



Fonte: Elaboração Própria com os Microdados da Prova São Paulo de 2008 e 2010.

Tabela 9 – Teste de Robustez – Efeito Fixo

	FE3- Matemática	FE3- Português
Usa sempre Caderno de Apoio	1.8574*** (0.6911)	2.2973*** (0.7869)
Placebo	-0.6468 (2.2503)	1.5634 (2.4634)
Controles: alunos e professores	sim	sim

Obs: Erros-padrão robustos entre parênteses. * p<0.10, ** p<0.05, *** p<0.01

Tabela 10 – Teste de Robustez para Caderno de Apoio de Matemática– Regressão Quantílica

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Intercepto	7.285*** (1.878)	20.917*** (1.485)	30.501*** (1.304)	39.533*** (1.150)	46.605*** (1.062)	54.631*** (1.176)	63.277*** (1.113)	73.774*** (1.280)	86.257*** (1.7600)
usa sempre Caderno de Apoio de Matemática	2.225 (1.378)	0.681 (0.945)	1.317 (0.895)	1.129 (0.822)	2.058*** (0.657)	2.270*** (0.792)	1.659** (0.817)	1.355 (1.006)	2.923*** (1.111)
Placebo	3.760 (4.327)	0.349 (2.678)	-0.632 (2.966)	-0.478 (2.509)	-1.025 (1.970)	-2.121 (2.355)	0.653 (3.903)	-1.703 (3.394)	0.566 (2.999)
Controles: alunos e professores	sim	sim	sim	sim	sim	sim	sim	sim	sim

Obs: Erros-padrão robustos entre parênteses. * p<0.10, ** p<0.05, *** p<0.01

Tabela 11 – Teste de Robustez para Caderno de Apoio de Leitura– Regressão Quantílica

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Intercepto	-4.776* (2.903)	13.036*** (2.045)	24.630*** (1.867)	32.249*** (1.643)	39.909*** (1.462)	47.715*** (1.807)	57.340*** (1.718)	65.507*** (1.723)	79.897*** (2.851)
Usa sempre Caderno de Apoio de Leitura	4.641* (2.442)	4.091*** (1.476)	3.791** (1.567)	2.928*** (1.075)	2.754** (1.1450)	1.498 (1.104)	1.738 (1.067)	3.6869*** (1.362)	1.711 (2.371)
Placebo	2.109 (4.466)	-1.621 (4.023)	-3.336 (3.649)	-1.647 (3.876)	-1.473 (5.070)	4.422 (5.730)	3.615 (4.042)	1.271 (3.124)	-6.047 (5.903)
Controles: alunos e professores	sim	sim	sim	sim	sim	sim	sim	sim	sim

Obs: Erros-padrão robustos entre parênteses. * p<0.10, ** p<0.05, *** p<0.01

5. Conclusão

O objetivo deste trabalho é estimar o impacto de uma nova política educacional implementada em 2010 pela Secretaria Municipal de Educação de São Paulo, o Caderno de Apoio e Aprendizagem, nas proficiências dos alunos da 4ª série. O intuito desse novo instrumento pedagógico é auxiliar os professores a seguirem as orientações curriculares da cada série e a cumprirem o cronograma de ensino previsto. As evidências apontadas pela literatura sobre *textbooks* é de que tais materiais impactam positivamente nos indicadores educacionais.

Para atingir esse objetivo, primeiramente estimamos por Primeiras Diferenças o efeito do Caderno de Apoio de Matemática e de Leitura nas suas respectivas proficiências para um painel de alunos. No entanto, uma das hipóteses do trabalho é que esse efeito é heterogêneo ao longo da distribuição de conhecimento acumulado entre a 2ª e a 4ª série. Por isso, também estimamos o modelo por Regressão Quantílica para a mesma especificação utilizada depois de tirada a Primeira Diferença. Os resultados evidenciam que o Caderno de Apoio apresenta um efeito médio positivo e significativo tanto na proficiência de matemática, como para de leitura. Já os resultados para os quantis mostram que o impacto do Caderno de Apoio de Matemática na proficiência é positivo e significativo somente a partir do 5º quantil, e apresenta um maior impacto no 9º quantil, isto é, entre os alunos com maior habilidade. Diferentemente, o Caderno de Apoio de Leitura apresenta um impacto na nota considerável nos quantis inferiores, com destaque para o 1º quantil. Tal diferença pode ser explicada pelo fato do aprendizado em matemática estar mais associado aos conhecimentos adquiridos na escola, enquanto o aprendizado em leitura está também associado ao hábito de leitura, fortemente influenciado por características familiares. Por fim, foi realizado um teste de robustez (explorando uma falseabilidade) que sinaliza que não há auto-seleção dos alunos.

Este trabalho inova por usar uma estrutura em painel de alunos, e também por o Caderno de Apoio poder ser considerado um tratamento homogêneo, tanto pelo fato de ser uma política educacional de uma única rede, como pelo fato de que os alunos que foram expostos ao Caderno de Apoio no mínimo três vezes na semana apresentam maior probabilidade de serem apresentados para o mesmo conteúdo. Isso reforça a necessidade do Caderno Apoio ser usado como principal material do professor em sala de aula, isto é, mais de três vezes por semana.

Ainda que os resultados aqui encontrados sejam potencialmente relevantes, testes de robustez adicionais são necessários para averiguar a validade dos resultados. Por exemplo, um teste de falseabilidade da auto seleção do professor em relação à adoção do Caderno Apoio só será possível após a divulgação dos dados da Prova São Paulo, quando poderemos reestimar o atual modelo usando uma variável “placebo” igual a 1 se o professor adotou o Caderno de Apoio somente em 2011 ou 2012, e 0 caso contrário.

Uma extensão final deste trabalho deve contar com duas análises adicionais já mencionadas pela literatura. A primeira, que já foi realizada em alguns estudos, seria a de custo efetividade que será factível de ser realizada uma vez que tenhamos acesso às informações sobre os Custos do Caderno de Apoio junto à Secretaria Municipal de Educação de São Paulo. A outra, ainda não explorada na literatura, consisti na investigação sobre os mecanismos pelos quais o Caderno de Apoio impacta positivamente na nota, isto é, quais seriam as modificações que a implementação do Caderno de Apoio inseriu no cotidiano de ensino e de aprendizado que resultou em melhoria da proficiência.

Bibliografia

Albernaz, A.; Ferreira, F. H. G.; Franco, C. 2002. Qualidade e Equidade na Educação Fundamental Brasileiro. *Pesquisa e Planejamento Econômico*, v. 33 No .3.

Angrist, J.; Pischke, J. (2009) *Mostly Hamrless Econometrics*. 1 ed. Princeton University Press.

Carnoy, M.; Loeb, S. 2002. Does external accountability affect student outcomes? A Cross-State Analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305–331.

Fuller, Bruce and Prema Clarke. 1994. “Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classrooms, Tools, Rules and Pedagogy”, *Review of Educational Research*, 64(1), 119-157.

Frölich, Melly (2010): Estimation of quantile treatment effects with STATA, *Stata Journal*, 10 (3), 423-457.

Fuller, B. Raising school quality in developing countries. *World Bank Discussion Papers*, Washington-DC, 1986.

Glewwe, P. Kremer M, Moulin S. Many Children Left Behind? Textbooks and Test Scores in Kenya. *NBER Working Paper Series WP 13300*, 2007.

Heyneman, Stephen P., Joseph P. Farrell and Manuel A. Sepulveda-Stuardo. 1978. Textbooks and Achievement: What We Know.” *World Bank Staff Working Paper No. 298*.

Jacob, B. A., 2005. Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools, *Journal of Public Economics*, Elsevier, vol. 89(5-6), pages 761-796, June.

Jamison, D. T.; Searle, B.; Galda, K.; Heyneman, S. P. Improving elementary mathematics education in Nicaragua: An experimental study of the impact of textbooks and radio on achievement. *Journal of Educational Psychology*, Vol 73(4), Aug 1981, 556-567.

Koenker, R. and Bassett, G. 1978. Regression quantiles. *Econometrica*, 46(1):33-50.

Leme, M.C.S, Louzano, P., Ponczek, V., Souza, A. P. The Impact of Structured Teaching Methods on the Quality of Education. *Economics of Education Review*, 2012.

Lockheed, M. Hanushek, E.A. Improving the Efficiency of Education in Developing Countries: Review of the Evidence. *Discussion Paper Education and Training Series*, The World Bank, 1987.

Rivkin, S. G., E. A. Hanushek, AND J. F. Kain, 2005. Variable Definitions, Data, and Programs for Teachers, Students, and Academic Achievement. *Econometrica Supplementary Material*, 73, 2.

Rubin, D. 1977. Assignment to a Treatment group on the basis of a Covariate. *Journal of Educational Statistics*. n. 2, p.1-26.

Pritchett, Lant, and Deon Filmer. 1999. “What Education Production Functions Really Show: A Positive Theory of Education Expenditures,” *Economics of Education Review*. 18(2):223-239.

Tan, J. P; Lane, G. L. 1999. Schooling Outcomes in Philippine Elementary Schools: Evaluation of the Impact of Four Experiments. *World Bank Economic Review*.

World Bank. Achieving world class education in Brazil: the next agenda, 2010. Disponível em: <<http://web.worldbank.org>>. Acesso em: jun. 2012.

Wooldridge, J. (2010) *Econometric Analysis of Cross Section and Panel Data*. 2 ed. Cambridge: MIT Press.

Apêndice

Tabela A.1 – Impacto do Caderno de Apoio de Matemática nos Quantis da Distribuição do Conhecimento Acumulado entre 2ª e 4ª série – Regressão Quantílica

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Intercepto	7.5647*** (1.8255)	20.8264*** (1.3335)	30.4112*** (1.1465)	39.3706*** (1.0945)	46.4189*** (0.8784)	54.6498*** (1.1037)	63.2562*** (1.0437)	73.7772*** (1.1418)	86.2570*** (1.4733)
usa sempre CA de Matemática	1.9770 (1.3310)	0.6930 (0.9867)	1.2819 (0.9310)	1.2290 (0.8497)	2.1381*** (0.7135)	2.4743*** (0.7091)	1.6223** (0.7949)	1.4316 (0.8839)	2.9231*** (1.0545)
tem muita lição	2.3493** (0.9398)	2.2274*** (0.6093)	1.6724*** (0.5287)	0.9144 (0.5984)	0.5724 (0.5237)	1.0517* (0.5963)	1.5283*** (0.5627)	1.3139** (0.5842)	0.9745 (0.7503)
pais ajudam bastante a fazer a lição	2.5563*** (0.9613)	2.5325*** (0.7551)	2.0259*** (0.6477)	1.4359* (0.7578)	1.2251* (0.6660)	0.7597 (0.6329)	-0.0111 (0.6000)	-0.9802 (0.7662)	-2.0606** (0.8849)
falta aula	-4.9969*** (1.0388)	-3.2597*** (0.8383)	-3.4244*** (0.7203)	-3.4803*** (0.8045)	-2.9790*** (0.6697)	-2.6917*** (0.6413)	-3.5502*** (0.7642)	-3.2209*** (0.8672)	-3.5272*** (1.0512)
professor corrige lição	0.6262 (1.4056)	-0.4088 (1.0080)	0.1203 (0.9453)	0.7064 (0.8469)	0.6382 (0.7411)	1.1465 (0.8729)	1.3601 (0.9984)	2.3082** (1.1303)	1.7545 (1.1072)
aprendo bastante	2.6958* (1.4058)	2.4953* (1.4017)	2.2396** (0.9458)	1.9068* (1.1375)	1.9432* (1.0586)	2.7539*** (1.0288)	1.2108 (1.1080)	-0.1741 (1.1850)	1.0089 (1.2568)
média de colegas cujos pais ajudam bastante a fazer a lição	15.3729*** (3.5308)	12.2309*** (2.6232)	8.0997*** (2.2913)	7.4722*** (2.5739)	5.2365** (2.1354)	4.8887** (2.0664)	4.0094 (2.4445)	4.0875* (2.3213)	6.1208** (2.9074)
média de colegas que entraram na 1ª série	-7.8932* (4.3224)	-3.1935 (3.2987)	-0.4073 (3.0271)	-0.9143 (3.1461)	0.2025 (2.2724)	-2.1452 (2.5352)	-3.1755 (2.6478)	-2.6846 (3.2946)	-1.6849 (3.4331)
professor frequentou mais que ES	1.0389 (1.2877)	2.0406** (0.9261)	1.4921 (0.9641)	1.4453 (0.9301)	1.5264** (0.7111)	0.7418 (0.7163)	0.9477 (0.8082)	1.4989 (0.9781)	1.0070 (1.2080)
professor mais de 10 anos de experiência	2.0966 (1.5126)	1.0349 (1.2734)	1.2441 (1.0689)	1.6413 (1.0307)	1.5801* (0.8919)	0.6106 (1.0399)	1.2939 (0.9138)	1.2535 (1.0688)	-0.3988 (1.5397)
perco mais da metade da aula para começar a aula	-2.7072** (1.2921)	-2.2237** (0.9753)	-2.5302*** (0.8334)	-1.9693** (0.9327)	-2.0697*** (0.7472)	-2.7149*** (0.8095)	-2.8810*** (0.9366)	-1.9333** (0.9858)	-2.9528*** (1.0212)
professor mora perto	0.0427 (1.1450)	0.1326 (0.7367)	0.1768 (0.6368)	0.0313 (0.8447)	0.1846 (0.5410)	-0.2594 (0.6033)	-0.7159 (0.7073)	-0.5040 (0.8741)	-0.9903 (0.8319)
professor acredita que mais de 80% da turma será aprovada	4.0374** (1.7271)	2.8576* (1.7343)	3.2305** (1.3844)	3.0741** (1.3380)	3.6772*** (1.2945)	4.6113*** (1.3075)	5.5202*** (1.6612)	3.7910* (1.9505)	4.9824*** (1.7044)
professor considera ter bastante violência	-0.9337 (1.7241)	-0.8702 (1.3306)	0.1763 (1.0232)	-0.2581 (1.2800)	-1.0299 (0.9295)	-2.2759** (1.0769)	-2.7230** (1.2842)	-2.6432 (1.6346)	-2.3741 (1.8173)
Observações	10141	10141	10141	10141	10141	10141	10141	10141	10141

Tabela A.2 – Impacto do Caderno de Apoio de Leitura nos Quantis da Distribuição do Conhecimento Acumulado entre 2ª e 4ª série – Regressão Quantílica

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Intercepto	-0.9955 (1.7373)	16.1648*** (1.1540)	25.4063*** (1.0428)	33.7260*** (1.0698)	41.4440*** (1.0834)	49.6474*** (1.3375)	58.5527*** (1.2958)	67.5767*** (1.2646)	81.7190*** (1.8949)
Usa sempre Caderno de Apoio de Leitura	3.9043*** (1.3456)	3.3006*** (0.9858)	3.2597*** (0.9380)	2.8971*** (0.7517)	2.0228*** (0.7821)	1.8933** (0.8537)	1.5731* (0.8048)	2.0241** (0.8983)	1.8040 (1.3231)
tem muita lição	0.8868 (1.0027)	1.4583* (0.7626)	1.5406** (0.6291)	1.4754** (0.6368)	1.4410*** (0.5311)	1.4442** (0.6742)	1.1290* (0.6073)	1.5681** (0.6253)	1.4740 (0.9097)
pais ajudam bastante a fazer a lição	0.2480 (1.0081)	1.5137** (0.7495)	1.4204** (0.6570)	1.1176* (0.6002)	1.3732** (0.5999)	0.7305 (0.7117)	0.4532 (0.6885)	0.7527 (0.7560)	0.3686 (1.0219)
falta aula	-3.5316*** (1.0122)	-3.4295*** (0.8371)	-2.5272*** (0.7730)	-2.3969*** (0.8778)	-2.2548*** (0.7889)	-2.3814*** (0.7297)	-2.1962** (0.8623)	-2.4381*** (0.8882)	-2.0790 (1.3036)
professor corrige lição	-2.1932 (1.4746)	0.2033 (0.9956)	0.4441 (0.7197)	0.3657 (0.7563)	-0.4421 (0.7297)	0.3974 (0.9681)	0.3259 (0.8807)	0.5881 (1.0007)	1.1715 (1.4054)
aprendo bastante	4.1157*** (1.4779)	3.5466*** (1.2586)	3.3606*** (0.8796)	3.4842*** (0.8148)	2.9792*** (0.8192)	3.5651*** (1.1505)	3.8813*** (1.1665)	4.0391*** (1.3503)	4.8152*** (1.6003)
média de colegas cujos pais ajudam bastante a fazer a lição	7.4799*** (2.8774)	5.2413** (2.0716)	5.4567*** (1.8835)	4.5503** (1.9265)	7.2337*** (1.9848)	7.5032*** (2.3597)	5.1728** (2.2114)	2.8483 (2.3610)	-0.7627 (3.6031)
média de colegas que entraram na 1ª série	-3.1697 (3.9571)	-5.2474** (2.4578)	-2.5746 (2.3826)	-4.3848* (2.2946)	-3.2848 (1.9979)	-0.6856 (2.5413)	-0.4458 (2.2647)	-1.4961 (2.9181)	-1.1600 (4.0423)
professor frequentou mais que ES	2.2266* (1.2752)	1.4654 (1.0598)	1.8175* (0.9434)	1.5066 (0.9181)	1.2822 (0.8988)	0.8451 (1.0271)	-0.2812 (1.0941)	-0.0207 (1.0123)	0.1438 (1.2719)
professor tem mais de 10 anos de experiência	-0.3145 (1.5761)	0.1955 (1.1121)	0.2329 (0.8128)	0.9114 (0.5653)	0.4116 (0.6997)	0.5285 (0.8821)	0.3249 (0.9127)	-0.2849 (0.9516)	1.1304 (1.4578)
perco mais da metade da aula para começar a aula	-2.7676** (1.3321)	-3.2353*** (1.0800)	-2.6813*** (0.9822)	-3.3768*** (0.8192)	-2.6279*** (0.9161)	-2.3730*** (0.8147)	-1.6108* (0.8333)	-1.1008 (0.9859)	-0.6218 (1.1577)
professor mora perto	-1.8538* (0.9922)	-1.5859* (0.8370)	-1.0345 (0.6520)	-0.8149 (0.6701)	-0.0173 (0.5918)	-0.1668 (0.7198)	-0.1465 (0.8035)	-0.4314 (0.7537)	1.0123 (1.0462)
professor acredita que mais de 80% da turma será aprovada	0.6354 (2.2459)	2.2004 (1.4477)	3.6207** (1.5831)	3.0233** (1.3006)	2.4615* (1.4015)	1.8568 (1.3599)	0.8258 (1.0804)	0.9620 (1.4652)	2.7295 (2.2294)
professor considera ter bastante violência na escola	0.8187 (1.7599)	-0.1019 (1.3851)	0.1399 (1.3499)	-0.0832 (1.0218)	-0.3139 (0.9912)	-1.2933 (1.0995)	-0.8704 (1.0556)	-1.3054 (1.4595)	-1.4954 (2.3937)
Observações	9635	9635	9635	9635	9635	9635	9635	9635	9635