

## OS DETERMINANTES DO APRENDIZADO COM DADOS DE UM PAINEL DE ESCOLAS DO SAEB

ANA MARIA DE PAIVA FRANCO \*  
NAERCIO AQUINO MENEZES FILHO †

### Resumo

Este trabalho analisa os impactos das características de alunos, professores, turmas, diretores e escolas sobre o desempenho em matemática de alunos da 4ª série, valendo-se de dois painéis de escolas construídos a partir de dados do SAEB 1997-2005. É a primeira vez que um painel de escolas com dados desagregados por alunos é utilizado para explorar este tema no caso brasileiro. O *background* familiar se mostrou muito importante para explicar desempenho, enquanto poucas variáveis de escolas apresentaram algum efeito. As análises apontam também para efeitos diferenciados das variáveis explicativas sobre o aprendizado dependendo da escola ser pública ou privada.

**Palavras-chave:** Educação; Painel de escolas; Aprendizado.

### Abstract

This work analyses the impacts of characteristics of students, teachers, classes, headmasters and schools on 4th grade student's outcomes in Mathematics with the use of two school's panel data sets built from SAEB 1997-2005. It is the first time that a panel of schools with disaggregated data at student level is used to explore this theme in the Brazilian case. The family background showed up to be very important for explaining achievement, whereas few school's variables had some effect. The analysis also indicate that explainable variables have differentiated impact on learning depending if the school is public or private.

**Keywords:** Education; School's panel; Learning.

**JEL classification:** I20, I21

**DOI:** <http://dx.doi.org/10.11606/1413-8050/ea120268>

---

\* Universidade Federal de Uberlândia. E-mail: ana.paiva@ufu.br

† INSPER. E-mail: naercioAMF@insper.org.br

## 1 Introdução

Desde que o INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, autarquia federal vinculada ao Ministério da Educação) passou a tornar público o *ranking* das escolas brasileiras, baseado no desempenho de alunos, a questão da qualidade da educação ganhou maior projeção na sociedade brasileira. Acalorou-se o debate entre os formuladores de políticas educacionais e os pesquisadores da área sobre as características das escolas que poderiam ser determinantes ou contribuir para explicar o diferencial de qualidade observado entre as escolas e, principalmente, quais delas seriam passíveis de intervenção política. Há muitas evidências sobre os impactos da qualidade da educação sobre a renda individual e crescimento econômico (Hanushek 1995, Mulligan 1999, Murnane et al. 2000, Lazear 2003).

A questão sobre os fatores determinantes da qualidade do ensino sempre foi alvo de investigação, e alguns resultados importantes vieram à tona nos últimos anos, em estudos oriundos de diversas partes do mundo. Este trabalho procura jogar mais luz na investigação sobre a qualidade no ensino para o caso brasileiro. Utilizando-se dados dos exames do SAEB e do Censo Escolar (ambos levantados pelo INEP), procura-se estimar os efeitos sobre a proficiência dos alunos de diversas características de suas famílias, professores, diretores e escolas, controlando-se pelos efeitos não observáveis específicos de cada escola. Para tal utilizaram-se dois conjuntos de dados em painel das escolas avaliadas nos exames do SAEB de 1997, 1999, 2001, 2003 e 2005.

A estimação dos coeficientes de regressão foi feita pelo método de mínimos quadrados ordinários ou generalizados, conforme o caso. Por se tratar de dados em painel, os modelos considerados foram o modelo de regressão agrupado, o modelo de Efeitos Fixos e o modelo de Efeitos Aleatórios. Os resultados dessas estimações são analisados à luz de outros estudos feitos para o caso nacional.

Além desta introdução, este trabalho está dividido da seguinte forma: a Seção 2 traz uma breve revisão da literatura sobre os determinantes da qualidade no ensino; a Seção 3 comenta as características dos dados do SAEB e Censo Escolar e traz as estatísticas descritivas das variáveis que entraram na análise; a Seção 4 comenta a metodologia empregada; a Seção 5 traz os resultados encontrados nas diversas análises e a Seção 6 traz os comentários e algumas conclusões que podem ser tiradas a partir do estudo realizado.

## 2 Referencial Teórico

Em 1966, foi realizada nos Estados Unidos uma grande pesquisa sobre educação promovida pelo *Office of Education* daquele país. O sociólogo James Coleman foi designado pelo Congresso dos Estados Unidos para levar a cabo um estudo em larga escala sobre as escolas americanas que recebiam alunos de diferentes raças no país. Acreditava-se, então, que intervenções sociais, financiadas pelo governo federal, pudessem corrigir as desigualdades entre brancos e negros, por exemplo. Os resultados a partir do *Coleman Report* surpreenderam aqueles que estavam preocupados em promover uma maior equidade no ensino. Concluiu-se que as diferenças nos recursos das escolas não eram muito relevantes para explicar a elevada desigualdade no aprendizado dos alunos de diferentes raças, e que o fator preponderante para explicar essa diferença seria o “*background*” familiar, em particular o status socioeconômico

dos alunos. A visão predominante passou a ser a de que apenas as famílias e os colegas de classe podiam afetar o desempenho escolar. Um resultado positivo do *Coleman Report* foi de servir de incentivo para que muitos se propusessem a prová-lo equivocadamente, isto é, a demonstrar que as escolas faziam sim diferença (Lee 2001, p. 38).

Hanushek (2002, 2005a), por exemplo, argumenta que no *Coleman Report*, bem como nos estudos posteriores que não conseguiram encontrar efeitos significativos de escolas e professores, teria havido uma confusão entre “mensurabilidade” e os verdadeiros efeitos das escolas sobre o aprendizado. O autor chama a atenção para dois resultados conclusivos em suas investigações: em primeiro lugar, há grandes e importantes diferenças entre os professores; em segundo lugar, essas diferenças não são captadas por medidas comuns de qualidade, tais como grau de qualificação e anos de experiência. Segundo sua definição de “qualidade”, os bons professores são aqueles que obtêm grandes ganhos de desempenho dos alunos de suas classes, sendo professores ruins, em contraposição, aqueles que não obtêm bons desempenhos (Hanushek 2002, 2005a). Hanushek (2005b, p. 17-18) argumenta ainda que as evidências internacionais apresentam em geral os mesmos resultados obtidos nos estudos para os Estados Unidos, ou seja, de que não há uma relação consistente e sistemática entre recursos extras nas escolas e o desempenho dos alunos.

No que diz respeito aos países em desenvolvimento, em geral há um ceticismo na literatura quanto à capacidade de estudos retrospectivos capturarem os efeitos sobre o aprendizado, seja de insumos escolares ou de programas não realizados nos moldes de um estudo experimental ou aleatório. Glewwe & Kremer (2006) fazem um levantamento dos diferentes tipos de estudos empíricos realizados nos países em desenvolvimento sobre a influência que características de professores e escolas exercem sobre o aprendizado nas escolas primárias e secundárias. Dentre eles está o de Hanushek & Pace (1995), que revê 96 estudos com dados retrospectivos sobre os impactos, no desempenho dos alunos, dos seguintes insumos educacionais: razão aluno por professor, experiência do professor, salário do professor, gastos por aluno e facilidades de infraestrutura. Hanushek conclui que, à exceção de “facilidades físicas”, os recursos mensurados não estão sistematicamente relacionados com o desempenho dos alunos nos países em desenvolvimento. Os resultados resumidos de sua análise podem ser vistos na Tabela 1.

**Tabela 1:** Resultados de 96 estudos dos efeitos estimados de recursos escolares sobre a aprendizagem (países em desenvolvimento)

Insumos	Número de Estudos	Estat. Significante		Estat. Insignificante
		Positivo	Negativo	
Razão aluno-professor	30	8	8	14
Educação do professor	63	35	2	26
Experiência do professor	46	16	2	28
Salário do professor	13	4	2	7
Gastos por aluno	12	6	0	6
Facilidades	34	22	3	9

Fonte: Hanushek & Pace (1995, p. 230).

Porém, há razões para se acreditar que, à exceção da “taxa aluno-professor”, os outros insumos exerçam um efeito positivo sobre a aprendizagem, pois é muito pequena a probabilidade de que tantos trabalhos apresentem estima-

tivas positivas quando o verdadeiro parâmetro é zero ou negativo (Kremer 1995). Outra interpretação possível é que talvez os estudos não estejam medindo os mesmos parâmetros, ou as estimativas sejam viesadas. Nesse caso, dependendo da estimativa, os parâmetros estimados podem apresentar viés positivo ou negativo devido ao problema de variável omitida, por exemplo (Glewwe & Kremer 2006, p. 30).

Em todas as análises com dados retrospectivos analisados por Glewwe & Kremer (2006), a maioria das variáveis de escola e professores não apresenta parâmetros significativamente diferentes de zero. Os autores alegam, porém, que isso pode decorrer do tamanho pequeno das amostras e de uma elevada correlação entre muitas das variáveis, além de grandes diferenças nos ambientes socioeconômicos dos países analisados e da variação na disponibilidade de algumas informações. Concluem que não existem resultados generalizáveis no que diz respeito a quais variáveis de escola e de professor elevam o aprendizado em países em desenvolvimento. Ressaltam ainda que, os principais problemas de estimação que esse tipo de análise enfrenta são a omissão de características das escolas e professores, características não observadas das crianças e suas famílias, que são correlacionadas com características observáveis de escolas e professores, e erros de medida nas variáveis de escola e professores (Ibid., p.34).

Na literatura nacional sobre educação, a questão da qualidade do ensino e da contribuição das escolas e professores para o aprendizado dos alunos também tem sido bastante explorada<sup>1</sup>. Porém, faz-se necessário continuar o esforço no sentido de avançar o debate sobre como melhorar os gastos em educação, para que estes não apenas aumentem sua participação no total dos gastos públicos, mas efetivamente promovam a melhoria da qualidade e equidade do ensino no país.

Partindo-se do pressuposto de que insumos escolares contribuem de alguma maneira para o aprendizado (já que não existe um consenso nem na literatura internacional e nem na nacional sobre esse ponto), e utilizando-se o método de Efeitos Fixos (que melhor pode minorar os inúmeros problemas que cabem no termo de “endogeneidade” nas estimações com dados retrospectivos), construiu-se um painel de escolas para um período de cinco anos de avaliações no SAEB e Censo Escolar e buscou-se avaliar o impacto sobre o desempenho dos alunos de uma série de insumos escolares. Esses insumos vão desde características dos professores, turmas, diretores e atributos físicos da escola até os programas sociais dos quais ela participa. Pelo fato de contar com um painel mais longo de dados do SAEB até aqui estudados, mais informações que incluem as características socioeconômicas do corpo discente, e de se realizar estimações com as observações individuais de cada aluno em cada escola, e não apenas com as médias por escolas, acredita-se que se este trabalho traz novos ingredientes e contribuições para a análise do caso nacional.

### 3 Dados

Os dados utilizados nesta pesquisa provêm do Sistema Nacional de Avaliação da Educação Básica (SAEB) e também do Censo Escolar. Criado em 1988, o SAEB foi desenvolvido pelo Instituto Nacional de Estudos e Pesquisas Edu-

---

<sup>1</sup>Ver Felício & Fernandes (2006), Menezes-Filho (2007), Biondi & Felício (2007), Franco et al. (2007), Alves & Soares (2007).

cacionais Anísio Teixeira (INEP), na sua Diretoria de Avaliação da Educação Básica (Daeb), com dados sobre alunos, professores, diretores de escolas públicas e privadas em todo o Brasil. De natureza amostral, o SAEB foi aplicado a cada dois anos desde 1993 para avaliar o desempenho dos alunos brasileiros da 4ª e da 8ª séries do ensino fundamental e da 3ª série do ensino médio, nas disciplinas de língua portuguesa (Foco: leitura) e matemática (Foco: resolução de problemas). Foram realizados vários levantamentos desde 1993, que sofreram aperfeiçoamentos em cada aplicação, tanto do ponto de vista metodológico como nos procedimentos, operacionalização e abrangência<sup>2</sup>.

O SAEB é baseado em uma amostra probabilística complexa, estratificada por estado e por dependência administrativa das escolas (privadas, públicas estaduais e públicas municipais). As unidades primárias de amostragem são as escolas e as secundárias são as turmas. Quando uma turma de uma escola é selecionada, todos os alunos dessa turma entram na amostra do SAEB, sendo que a metade dos alunos é avaliada em leitura e a outra metade em matemática (Franco et al. 2007).

Nos diversos anos em que o SAEB é realizado, algumas escolas se repetem nas amostras. Isso permite que se explore a metodologia para análise de dados em painel, de forma a buscar maior precisão nos parâmetros de interesse. As estimativas dos efeitos de variáveis regressoras observáveis apenas com dados referentes a *cross sections*, como é comum no caso dos estudos sobre insumos de educação no Brasil, não permitem o controle de características não observáveis nas escolas. Isso pode fazer com que essas estimativas apresentem viés de variável omitida se essas variáveis estiverem correlacionadas com os demais regressores.

Este trabalho utilizará os dados do SAEB referentes às escolas da rede pública e privada onde foram aplicadas as provas de matemática para a 4ª série do ensino fundamental nos anos de 1997, 1999, 2001, 2003 e 2005, em todas as unidades da federação. A escolha da proficiência em matemática como variável dependente é usual na literatura e reflete um aprendizado tipicamente de natureza escolar, o que não ocorre com o conhecimento da linguagem, que se desenvolve em vários ambientes frequentados pelas crianças (Franco et al. 2007).

Como uma das maiores dificuldades em se construir um painel de escolas com os dados do SAEB é compatibilizar as respostas dos dicionários (alunos, professores, diretores, turmas e escolas) nos diferentes anos sem se abrir mão

---

<sup>2</sup>Os resultados das diversas avaliações podem ser comparados, pois a partir de 1995 foi introduzido o uso da Teoria de Resposta ao Item (TRI) para a construção de instrumentos, a atribuição de scores e a análise, para permitir a comparação dos dados. Os resultados obtidos a partir da TRI são independentes de grupos e não são afetados pela dificuldade dos testes. A comparabilidade é garantida por meio da inclusão de itens comuns às avaliações de edições anteriores, o que recebe o nome de “equalização de grupos não equivalentes com itens comuns”. Utilizam-se matrizes de referência como instrumentos, que servem de orientação para a construção e seleção dos itens de avaliação (INEP 2009a). A partir de 2005 o SAEB foi reestruturado e passou a ser composto por duas avaliações: Avaliação Nacional da Educação Básica (Aneb) e Avaliação Nacional do Rendimento Escolar (Anresc), conhecida como Prova Brasil. Em 2013, também foi encampada pelo SAEB a Avaliação Nacional da Alfabetização (ANA). A Aneb deu continuidade aos procedimentos amostrais (atendendo aos critérios estatísticos de no mínimo 10 estudantes por turma), das redes públicas e privadas, com o mesmo foco do SAEB. Já a Anresc (Prova Brasil) passou a ser uma avaliação de caráter censitário das escolas públicas que atendessem a critérios de quantidade mínima de estudantes na série avaliada, permitindo gerar resultados por escola. Em 2013, passou a fazer parte do SAEB também a Avaliação Nacional da Alfabetização (ANA), prevista no Pacto Nacional pela Alfabetização na Idade Certa – PNAIC (INEP 2009b).

de variáveis que são importantes na função de produção escolar, serão reportados também os resultados obtidos com um painel mais curto de escolas, referente aos anos 1999, 2001, 2003 e 2005. Nesse caso, há maior proximidade nas perguntas, o que permite que sejam explorados os efeitos de mais variáveis explicativas, e onde também é maior o número de escolas que se repetem<sup>3</sup>.

Para completar o rol de informações referentes às escolas que entraram na análise de painel, foram incluídos os dados disponíveis no Censo Escolar (INEP), que contém informações sobre algumas características observáveis das escolas, seus alunos e professores. O Censo Escolar é realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP/MEC), em parceria com os governos estaduais (secretarias estaduais de educação) e prefeituras municipais. Ele contém informações sobre características das escolas, como insumos de produção e infraestrutura, composição do corpo docente em termos de modalidade de ensino, níveis de instrução dos professores, taxas de reprovação, aprovação e abandono, número de matrículas e turmas nos diferentes turnos entre outras.

### 3.1 Variáveis e estatísticas descritivas

Como o SAEB tem caráter amostral, muitas escolas não se repetem em todos os anos realizados. Dessa forma, para se utilizar a metodologia de dados em painel e prosseguir nas investigações às quais este trabalho se propõe, a análise é feita utilizando-se quatro conjuntos de dados:

1. Painel não balanceado de escolas para os anos 1997, 1999, 2001, 2003 e 2005, separado para escolas públicas e privadas;
2. Painel não balanceado de escolas para os anos 1999, 2001, 2003 e 2005, separado para escolas públicas e privadas.

É importante ressaltar que, por se tratarem de sub amostras de escolas que se repetem nos levantamentos do SAEB, os pesos probabilísticos das unidades primárias de amostragem (escolas) se alteram em cada ano e, por esse motivo, não puderam ser considerados nas análises estatísticas e regressões<sup>4</sup>. Isso pode por em cheque a validade externa dos resultados encontrados neste trabalho, mas é uma limitação dos dados que só pode ser contornada de alguma forma com levantamentos mais recentes da Prova Brasil de natureza censitária (para escolas públicas).

Para trabalhar com um painel de escolas, foram excluídas aquelas que só apareciam uma vez nos cinco anos compreendidos no período 1997-2005 (42,64% do total da amostra). Foram consideradas apenas as escolas que apresentavam número de matrículas na 4ª série no Censo Escolar maior ou igual a dez, e no mínimo uma turma dessa série em cada escola, de forma a tentar minimizar o problema de erro que existe na computação de dados como o

---

<sup>3</sup>Por exemplo, no ano de 1997 não há um questionário específico para as turmas, o que inviabilizou conseguir variáveis passíveis de comparação com as dos demais anos, que descrevessem a quantidade de alunos presentes na sala, a quantidade de horas de aula por dia, se a turma ficou sem professor de matemática por algum período (mais de um mês) e se a turma teve mais de um professor de matemática no período.

<sup>4</sup>Tentou-se construir um peso fixo para cada escola, com base na média dos pesos ao longo do período analisado, mas essa estratégia não se mostrou adequada.

Censo Escolar<sup>5</sup>. Além disso, foram excluídas as escolas federais, que representavam menos de 0,4% das escolas e que apareciam apenas nos anos de 2003 e 2005. As Tabelas 2 e 3 trazem as distribuições de frequência de alunos e de escolas por ano, segundo o tipo de escola, em cada sub amostra utilizada neste trabalho (painéis 1997-2005 e 1999-2005, respectivamente).

**Tabela 2:** Frequências de alunos e escolas por ano em que aparecem no painel de escolas 1997 a 2005

Escolas Públicas			Escolas Privadas		
Ano	Alunos	Escolas	Ano	Alunos	Escolas
1997	1.053	44	1997	148	3
1999	4.761	699	1999	3.018	511
2001	12.884	980	2001	7.843	651
2003	11.337	738	2003	5.209	379
2005	10.766	742	2005	6.861	545
Total	40.801	3.203	Total	23.079	2.089

Fonte: Cálculos próprios a partir do SAEB 1997-2005.

**Tabela 3:** Frequências de alunos e escolas por ano em que aparecem no painel de escolas 1999 a 2005

Escolas Públicas			Escolas Privadas		
Ano	Alunos	Escolas	Ano	Alunos	Escolas
1999	3.691	569	1999	2.467	436
2001	10.780	870	2001	6.863	593
2003	9.891	669	2003	4.702	348
2005	8.811	634	2005	5.679	464
Total	33.173	2.742	Total	19.711	1.841

Fonte: Cálculos próprios a partir do SAEB 1997-2005.

As variáveis que entraram nas análises e suas respectivas médias e desvios-padrão, bem como o teste de diferenças de médias e proporções entre a amostra do SAEB como um todo e a amostra do painel 1997-2005, estão descritas nas Tabelas 4, 5, 6, 7 e 8 para alunos, professores, diretores, turmas e escolas, respectivamente. Essas estatísticas descritivas estão separadas pelo tipo de administração da escola, se privada ou pública (escolas estaduais e municipais)<sup>6</sup>.

A Tabela 4 mostra que as características dos alunos da amostra completa do SAEB são bastante similares as dos alunos da amostra de escolas do painel 1997-2005, embora difiram estatisticamente em alguns aspectos. Contudo, os tamanhos das amostras aqui comparadas são muito grandes, o que

<sup>5</sup>Para se ter uma ideia do índice de erro nos dados, quase 13% das escolas que apresentavam dados sobre o desempenho dos alunos de 4ª série no SAEB apresentavam número de matrículas igual a zero no Censo escolar. Ao cortar o banco e ficarmos apenas com as escolas cujo número de matrículas é maior ou igual a dez, foram excluídas do banco 17,79% de escolas que apareciam pelo menos uma vez no período 1997-2005. Isso reduziu bastante a amostra, por outro lado se mostrou imprescindível, caso contrário poderia comprometer os resultados das estimativas.

<sup>6</sup>As tabelas com as estatísticas descritivas referentes ao painel 1999-2005 foram omitidas por conta de espaço e estão disponíveis sob requerimento aos autores.

**Tabela 4:** Variáveis de alunos na amostra SAEB e painel 1997-2005

Variáveis de Alunos	Escolas públicas			Escolas privadas		
	Saeb	Painel	Diferença (P-valor)	Saeb	Painel	Diferença (P-valor)
Proficiência Média	172,839	173,282	0,048**	219,438	224,886	0,000***
Homens	0,500	0,500	0,255	0,506	0,511	0,159
Idade 9 anos	0,090	0,090	0,055*	0,177	0,176	0,675
Idade 10 anos	0,370	0,390	0,000***	0,618	0,642	0,000***
Idade 11 anos	0,230	0,220	0,512	0,150	0,143	0,007***
Idade 12 ou mais	0,310	0,290	0,000***	0,054	0,040	0,000***
Branco	0,370	0,360	0,732	0,509	0,522	0,001***
Pardo/ Mulato	0,410	0,420	0,067*	0,352	0,344	0,035**
Negro	0,130	0,130	0,903	0,048	0,041	0,000***
Amarelo/ Ind.	0,070	0,070	0,967	0,078	0,080	0,469
<i>Mãe do aluno:</i>						
Nunca estudou	0,080	0,070	0,000***	0,012	0,009	0,001***
1ª à 4ª EF	0,290	0,290	0,032**	0,065	0,050	0,000***
5ª à 8ª EF	0,190	0,190	0,420	0,090	0,077	0,000***
Ensino médio	0,130	0,140	0,011**	0,203	0,199	0,248
Ensino Superior	0,070	0,070	0,000***	0,348	0,396	0,000***
Aluno não sabe	0,280	0,250	0,000***	0,304	0,269	0,000***
Aluno tem computador	0,100	0,100	0,033**	0,485	0,556	0,000***
Mora com o pai e mãe	0,660	0,640	0,000***	0,742	0,741	0,731
Observações:	129.284	40.801		52.783	23.079	

Fonte: Cálculos próprios a partir do SAEB (INEP/ MEC) 1997-2005.

Notas: em 2005 trata-se de computador com internet. Desvios-padrão em parênteses.

Testes de diferença de médias estatisticamente significantes com níveis de significância iguais a: \*\*\*1%, \*\* 5%, \* 10%.

faz com que os erros-padrão sejam artificialmente comprimidos e os testes de diferenças de médias e proporções apontem para diferenças significativas mesmo quando as características dos grupos não são economicamente relevantes, como nos casos das categorias de idade (há uma proporção maior de alunos na faixa de idade correta para a série na amostra painel de escolas públicas e privadas), proporção de alunos pardos e negros (maior proporção de pardos nas escolas públicas do painel e menor proporção de pardos e negros na amostra painel de escolas privadas), brancos (maior proporção na amostra painel de escolas privadas) e dos que moram com os pais (menor proporção na amostra painel de escolas públicas).

No que diz respeito à proficiência em matemática, nota-se que os alunos do painel apresentam desempenho médio melhor do que a amostra do SAEB – diferenças de médias estatisticamente significantes a menos de 1% de nível de significância – tanto para alunos da escola pública quanto privada. A diferença de meio ponto (0,5) no caso dos alunos das escolas públicas e de cinco e meio (5,5) no caso das escolas particulares representa, respectivamente, 0,001e 0,12 desvios-padrão das distribuições de proficiência<sup>7</sup>. No mesmo sentido, as mães dos alunos das amostras do painel apresentam uma escolarização mais alta em média do que no caso das amostras SAEB. Por exemplo, nas escolas públicas a proporção de mães que nunca estudaram é um ponto percentual menor entre os alunos do painel *vis a vis* aos alunos da amostra SAEB completa. Já no caso das escolas privadas a proporção de mães

<sup>7</sup>Desvios-padrão da distribuição de notas em matemática 4ª série: 39,55 escolas públicas, 46,77 escolas privadas.

com ensino superior entre os alunos do painel supera em cinco pontos percentuais a proporção observada na amostra SAEB.

Pelo descrito acima, os alunos que compõem o painel de escola tendem a apresentar desempenho médio e *background* familiar (captado aqui principalmente pela escolaridade da mãe) mais favoráveis que os da amostra Saeb completa<sup>8</sup>. Acredita-se, contudo, que essas diferenças não sejam economicamente significantes a ponto de comprometerem a validade externa dos resultados deste trabalho. Além do mais, estratégia de estimação da função de produção escolar por dados em painel, que controla pelos fatores não observáveis fixos no tempo que podem estar correlacionados com as variáveis explicativas do desempenho escolar, contorna o problema de potencial viés de variáveis omitidas correlacionadas com a seleção.

**Tabela 5:** Variáveis de professores na amostra SAEB e painel 1997-2005

Variáveis de Professores	Escolas públicas			Escolas privadas		
	Saeb	Painel	Diferença (P-valor)	Saeb	Painel	Diferença (P-valor)
<i>Sexo professor:</i>						
Homem	0,111	0,104	0,269	0,061	0,053	0,217
<i>Idade professor:</i>						
Até 29 anos <sup>(a)</sup>	0,256	0,256	0,000***	0,339	0,288	0,000***
De 30 a 40 anos <sup>(b)</sup>	0,339	0,339	0,978	0,370	0,401	0,012**
40 ou mais anos <sup>(c)</sup>	0,346	0,393	0,000***	0,244	0,264	0,078*
<i>Escolaridade:</i>						
Até o ensino médio	0,525	0,476	0,000***	0,400	0,312	0,000***
Tem ensino superior	0,408	0,460	0,000***	0,539	0,626	0,000***
<i>Experiência:</i>						
15 anos ou menos <sup>(d)</sup>	0,583	0,550	0,001***	0,663	0,611	0,000***
Mais de 15 anos <sup>(e)</sup>	0,355	0,392	0,000***	0,287	0,338	0,000***
Fez capacitação <sup>(f)</sup>	0,776	0,791	0,059*	0,798	0,828	0,003***
Observações	9.850	3.203		5.151	2.089	

Fonte: Cálculos próprios a partir do SAEB (INEP/ MEC) 1997-2005.

Notas: (1) não há informação de cor do professor para 1997 e, portanto, essa variável não entrará nas regressões do painel 1997-2005; (a) 1997 e 1999 incluem 30 anos; (b) 1997 e 1999 incluem 40 anos; (c) 1997 e 1999 a partir de 41 anos; (d) 2001 há até 14 anos; (e) 2001 há mais de 14 anos; (f) nos últimos 2 anos/no próprio ano para 1997.

Testes de diferença de médias estatisticamente significantes com níveis de significância iguais a: \*\*\* 1%, \*\* 5%, \* 10%.

A Tabela 5 mostra que as médias e proporções das características dos professores nas amostras painel e Saeb são bastante similares. Em termos estatísticos, porém, as escolas que compõem o painel possuem professores mais escolarizados e experientes em média tanto no caso das escolas públicas quanto no caso das privadas. Por exemplo, a proporção de professores com ensino superior nas escolas públicas do painel é seis pontos percentuais superior àquela da amostra SAEB, e no caso das escolas privadas essa diferença é de quase nove pontos percentuais (em termos de desvios-padrão das distribuições essas diferenças equivalem a 0,12 e 0,18, respectivamente). Quanto à experiência do professor, há uma diferença de quatro pontos percentuais (0,08 desvios-

<sup>8</sup>Observa-se o mesmo padrão na comparação das características dos alunos das escolas do painel 1999-2005 e Saeb.

padrão) na proporção de professores mais experientes na amostra do painel relativamente à amostra SAEB no caso das escolas públicas e nas escolas privadas essa diferença é de cinco pontos percentuais (0,10 desvios-padrão). Portanto, assim como se observou em relação às características dos alunos, as escolas do painel tendem a apresentar algumas características dos professores mais favoráveis do que a amostra SAEB, embora a significância econômica dessas diferenças seja questionável<sup>9</sup>.

**Tabela 6:** Variáveis de diretores na amostra SAEB e painel 1997-2005

Variáveis de Diretores	Escolas públicas			Escolas privadas		
	Saeb	Painel	Diferença (P-valor)	Saeb	Painel	Diferença (P-valor)
<i>Exerce a função de diretor há:</i>						
Menos de 5 anos	0,551	0,570	0,065*	0,234	0,215	0,0959*
Entre 5 e 10 anos	0,283	0,284	0,855	0,298	0,272	0,0280**
Mais de 10 anos	0,131	0,128	0,634	0,447	0,491	0,0000***
<i>É diretor dessa escola há:</i>						
Menos de 5 anos	0,682	0,713	0,001***	0,348	0,339	0,4860
Entre 5 e 10 anos	0,214	0,202	0,145	0,302	0,292	0,3660
Mais de 10 anos	0,069	0,069	0,979	0,328	0,349	0,0910*
<i>Assumiu o cargo por:</i>						
Seleção/eleição	0,452	0,477	0,014**	0,563	0,562	0,9910
Indicação/outras	0,501	0,498	0,711	0,181	0,202	0,0460**
<i>Projeto pedagógico:</i>						
Secretaria Educação	0,220	0,186	0,000***	0,103	0,090	0,1050
Diretor e Professores	0,603	0,655	0,000***	0,813	0,829	0,1100
Não houve no ano	0,121	0,116	0,460	0,041	0,035	0,2290
Há professores faltosos	0,183	0,217	0,000***	0,055	0,054	0,7720
Observações	9.850	3.203		5.151	2.089	

Fonte: Cálculos próprios a partir do SAEB (INEP/ MEC) 1997-2005.

No que tange às características dos diretores, a Tabela 6 mostra que na maioria dos casos não há diferenças economicamente relevantes entre as amostras painel e SAEB para escolas públicas e privadas, embora algumas diferenças tenham apresentado relevância estatística. Por exemplo, no que diz respeito às variáveis que captam a forma pela qual o diretor assumiu o cargo, há uma maior proporção de diretores que assumiram o cargo por exame de seleção e/ou eleição nas escolas públicas do painel (diferença de dois pontos percentuais) e uma maior participação daqueles que assumiram por indicação ou outras formas nas escolas privadas do painel (diferença de dois pontos percentuais), *vis a vis* às amostras do SAEB – diferenças estatisticamente significantes a 5% de nível de significância. Ou seja, há um indício economicamente não muito relevante de que na amostra painel de escolas públicas o processo de escolha dos diretores seja mais democrático que na amostra SAEB, e o contrário é observado na comparação de escolas privadas<sup>10</sup>.

<sup>9</sup>O mesmo padrão se observa no caso da comparação das características dos professores no painel 1999-2005 e Saeb.

<sup>10</sup>Padrões muito similares são observados nas comparações relativas ao painel 1999-2005. Assim, as escolas públicas do painel tendem a apresentar processos mais democráticos que as da amostra Saeb como um todo, enquanto que nas escolas privadas vale o contrário: na amostra painel prevalecem processo de escolha do diretor por indicação ou outras formas. Outra diferença estatisticamente significativa entre as amostras painel e Saeb se dá na questão relativa à elaboração

**Tabela 7:** Variáveis de turmas na amostra SAEB e painel 1997-2005<sup>(a)</sup>

Variáveis de Turmas	Escolas públicas			Escolas privadas		
	Saeb	Painel	Diferença (P-valor)	Saeb	Painel	Diferença (P-valor)
Não teve prof. Mat.	0,046	0,045	0,814	0,047	0,046	0,835
Teve um prof. Mat.	0,797	0,779	0,045**	0,857	0,878	0,025**
2 ou mais profs. Mat.	0,157	0,176	0,019**	0,096	0,076	0,012**
Nenhum mês sem prof.	0,958	0,957	0,908	0,978	0,983	0,155
1 ou + meses s/ prof.	0,042	0,043	0,822	0,022	0,017	0,187
<i>Turno da turma:</i>						
Até 4 horas por dia	0,601	0,579	0,034**	0,447	0,369	0,000***
Entre 4 e 5 horas	0,368	0,392	0,022**	0,537	0,616	0,000***
Mais de 5 horas	0,031	0,030	0,707	0,016	0,015	0,779
<i>Tamanho médio da turma</i>	27,565	29,154	0,000***	20,465	23,918	0,000***
Observações	9.009	2.742		4.927	1.841	

Fonte: Cálculos próprios a partir do SAEB e Censo Escolar (INEP/MEC).

Notas: (a) Não há questionário de Turmas no SAEB para o ano 1997; Testes de diferença de médias estatisticamente significantes com níveis de significância iguais a: \*\*\* 1%, \*\* 5%, \* 10%.

A Tabela 7 mostra as características das turmas que entrarão nas análises deste trabalho e permite visualizar os casos em que há diferenças estatisticamente significantes entre as escolas do painel e da amostra SAEB. As diferenças encontradas não são, porém, muito relevantes do ponto de vista econômico e em geral apontam que na amostra painel de escolas públicas as turmas tendem a ter mais de um professor de matemática no ano e turnos maiores que na amostra SAEB, enquanto que nas escolas privadas da amostra painel prevalecem turmas com menos rotatividade de professores de matemática e turnos também maiores que na amostra SAEB. Por fim, as turmas tendem a ser um pouco maiores nas escolas do painel, tanto no caso das públicas quanto das privadas.

A Tabela 8 abaixo traz as variáveis de escola que entram neste estudo e as respectivas médias e proporções observadas na amostra painel e Saeb. Observam-se diferenças estatisticamente significantes entre as amostras praticamente em todas as variáveis. Em alguns casos, essas diferenças são bastante discrepantes também em termos de magnitude. Por exemplo: há uma forte sub-representação das escolas localizadas na região rural nas amostras de escolas do Painel, tanto públicas quanto privadas. No caso das escolas públicas do painel, há uma maior representação de escolas estaduais em detrimento das municipais quando se compara com a amostra SAEB. Além disso, em termos da existência de insumos como biblioteca, laboratórios e quadras e esportes, tanto as escolas públicas quanto as privadas que compõem o Painel são mais bem dotadas do que as escolas da amostra SAEB. Em termos de participação em programas como TV Escola do MEC e merenda, as escolas públicas do painel também apresentam uma maior frequência frente à amostra Saeb. Por fim, destaca-se o maior número de matrículas totais na 4ª série nas esco-

do projeto pedagógico da escola: nas escolas públicas do painel há uma participação da categoria “projeto pedagógico foi desenvolvido por professores e diretores” maior que nas escolas da amostra SAEB (0,66 e 0,60, respectivamente), e uma menor participação da categoria “projeto pedagógico desenvolvido pela secretaria da educação” (0,19 versus 0,22). Por fim, as escolas públicas do painel apresentam uma proporção maior de “problemas com professores faltosos” que as da amostra SAEB (0,22 versus 0,18, respectivamente).

las do painel *vis a vis* às amostras do SAEB, tanto no caso das escolas públicas (83 *versus* 58, respectivamente) quanto nas privadas (56 *versus* 37, respectivamente). O número de turmas na quarta série do período diurno e o número de professores no ensino fundamental também são maiores nas escolas do Painel relativamente às escolas da amostra geral do SAEB. Assim, aparentemente as escolas do painel em geral são mais bem servidas em termos de insumos observáveis e também maiores em média que a as da amostra Saeb.

**Tabela 8:** Variáveis de escolas na amostra SAEB e painel 1997-2005

Variáveis de Escolas	Escolas públicas			Escolas privadas		
	Saeb	Painel	Diferença (P-valor)	Saeb	Painel	Diferença (P-valor)
Localização região rural	0,273	0,129	0,000***	0,010	0,005	0,0525*
<i>Administração:</i>						
Estadual	0,425	0,503	0,000***			
Municipal	0,575	0,497	0,000***			
Escola tem biblioteca	0,424	0,517	0,000***	0,817	0,896	0,0000***
<i>Laboratórios:</i>						
De Informática	0,125	0,164	0,000***	0,598	0,739	0,0000***
De Ciências	0,082	0,114	0,000***	0,392	0,531	0,0000***
<i>Tem quadra de esportes</i>	0,368	0,470	0,000***	0,660	0,797	0,0000***
<i>Programas sociais:</i>						
TV Escola (MEC)	0,510	0,572	0,000***	0,035	0,035	0,9350
Merenda Escolar	0,961	0,983	0,000***	0,067	0,068	0,9250
<i>Números na escola de:</i>						
Professores ensino fundamental	19,935	23,939	0,000***	18,744	24,276	0,0000***
Turmas 4ª série período diurno	1,810	2,468	0,000***	1,519	2,027	0,0000***
Turmas 4ª série noturno	0,340	0,156	0,000***	0,176	0,013	0,0000***
Matrículas totais na 4a série	58,190	82,993	0,000***	37,367	55,980	0,0000***
Total de escolas	9.009	2.742		5.151	2.089	

Fonte: Cálculos próprios a partir do Censo Escolar (INEP/MEC).

Notas: Testes de diferença de médias estatisticamente significantes com níveis de significância iguais a: \*\*\* 1%, \*\* 5%, \* 10%.

Como se pode verificar nas análises desta seção, observam-se algumas diferenças econômica e estatisticamente significantes entre as amostras Saeb e as amostras de escolas que compõem os painéis deste estudo<sup>11</sup>. Pode-se dizer que, em geral, as escolas do painel apresentam vantagens comparativas em termos de características socioeconômicas dos alunos, professores, turmas e escolas, o que pode ensejar questionamentos quanto à validade externa dos resultados encontrados nas análises econométricas. Contudo, mesmo havendo algumas diferenças entre as amostras painel e Saeb que se mostraram mais importantes em termos de magnitude, acredita-se que o potencial viés de estimação decorrente da seleção não aleatória da amostra painel possa ser contornado pela estratégia de estimação de dados em painel, como será explicado na seção que se segue.

<sup>11</sup>Novamente é importante lembrar que o erro-padrão das estimativas converge para zero devido ao grande número de observações nas amostras, o que leva à rejeição das hipóteses de igualdade das médias e proporções dos grupos a níveis de significância muito baixos.

#### 4 Metodologia

No presente trabalho, a questão é investigar o quanto algumas características observáveis das escolas e professores podem contribuir para o aprendizado dos alunos. É importante haver professores com elevado grau de instrução ou experiência? A forma com que é selecionado o diretor da escola é importante para a qualidade do ensino? Em geral, os alunos se saem melhor se estudam em turmas grandes, médias ou pequenas? Um dos grandes desafios em trabalhos de economia aplicada envolve responder o que aconteceria se uma situação diferente da qual foi ou pode ser observada se passasse com a unidade de observação – o chamado “contrafactual”. Isso porque observamos apenas se as escolas apresentam ou não determinadas características, sem poder observar sua situação se elas tivessem características diferentes<sup>12</sup>.

A estratégia de identificação que se utilizará no presente estudo é explorar a natureza dos dados em painel para se controlar os efeitos específicos não observados das escolas fixos no tempo que, em uma estrutura de dados de corte transversal e estimação por mínimos quadrados ordinários, poderiam viesar as estimativas. Pretende-se examinar os impactos que uma série de insumos e características escolares (variáveis causais) exerce sobre o aprendizado em matemática (variável dependente). É coerente supor que existam características não observáveis das escolas que também influenciam o aprendizado e que podem estar correlacionadas com os insumos e características escolares aqui analisados. Além disso, na literatura sobre o tema, é sabido que as características socioeconômicas dos alunos contribuem fortemente para explicar o desempenho e, por esse motivo, estas serão incluídas nas análises como variáveis de controle<sup>13</sup>.

Este trabalho emprega as metodologias de mínimos quadrados agrupados (MQO), efeitos fixos (EF) e efeitos aleatórios (EA) para estimar a seguinte equação que modela o desempenho dos alunos na prova de matemática do SAEB:

$$Y_{ijt} = \alpha_0 + A_{ijt}\beta_1 + P_{ijt}\beta_2 + D_{ijt}\beta_3 + T_{ijt}\beta_4 + E_{ijt}\beta_5 + c_j + u_{ijt} \quad (1)$$

em que:

$Y_{ijt}$  é a nota do aluno  $i$  na escola  $j$  no tempo  $t$ ;

$\alpha_0$  é uma constante;

$A_{ijt}$  vetor de características do aluno  $i$  na escola  $j$  no tempo  $t$ ;

$P_{ijt}$  vetor de características do professor de matemática do aluno  $i$  na escola  $j$  no tempo  $t$ ;

$D_{ijt}$  vetor de características do diretor do aluno  $i$  na escola  $j$  no tempo  $t$ ;

$T_{ijt}$  vetor de características da turma do aluno  $i$  na escola  $j$  no tempo  $t$ ;

$E_{ijt}$  vetor de características de infraestrutura da escola do aluno  $i$  no tempo  $t$ ;

<sup>12</sup>Essa é a dificuldade em lidar com situações “contrafactuais”. Tudo o que se pode observar é o resultado de apenas um dos cenários, para cada unidade de observação. Em muitos campos da ciência, especialmente nas ciências médicas, a visão predominante é de que a melhor evidência sobre contrafactuais é gerada por experimentos aleatórios, pois atribuir de forma aleatória um tratamento garante que os resultados obtidos no grupo de controle realmente capturem o contrafactual para um grupo de tratados (Angrist & Krueger 1999, p. 1283).

<sup>13</sup>É importante atentar para a distinção entre variáveis causais e variáveis de controle. A estratégia de identificação consiste na combinação de uma fonte de variação na variável que se quer analisar o efeito causal e o uso de técnicas econométricas para explorar essa informação, mantendo tudo o mais constante. Isto é, variáveis de controle são muitas vezes incluídas nas regressões para que os seus efeitos não se confundam com os efeitos da variável causal, que são o objeto de interesse primordial.

$c_j$  efeito específico não observado da escola;

$u_{ijt}$  erro idiossincrático.

O vetor de parâmetros de interesse é  $\beta$ , tal que  $\beta' = (\beta'_1, \beta'_2, \beta'_3, \beta'_4, \beta'_5)$  reporta as relações condicionais dos insumos escolares sobre o desempenho dos alunos. A Equação (1) é a função de produção de educação, como conhecida na literatura.

## 5 Resultados

Na Tabela 9, encontram-se os resultados das estimações realizadas por MQO agrupados e efeitos fixos da Equação (1) utilizando-se o painel de escolas mais longo para escolas públicas e privadas, respectivamente. Tanto nas estimações de MQO agrupados quanto nas estimações de efeitos fixos foram utilizadas as observações de cada aluno nas regressões. As estimações de MQO foram reportadas para serem comparadas com as obtidas por EF. Porém, como mencionado na seção anterior, o método de MQO agrupados é enviesado na presença de variável omitida que esteja correlacionada com o erro. As estimações por efeitos aleatórios não foram reportadas, pois os resultados do teste de Hausman levaram a que se rejeitasse a hipótese de que o modelo de efeitos aleatórios é o mais adequado<sup>14</sup>.

A Tabela 10 reporta as estimações realizadas com o painel mais curto para escolas públicas e privadas.

### 5.1 Resultados obtidos para as características dos alunos

Em todas as análises das Tabelas 9 e 10, os resultados mostram que os alunos do sexo masculino têm desempenho melhor nos testes de matemática do SAEB para a 4ª série do que as meninas, pois os parâmetros estimados apresentam sinal positivo e significativo a 1%. Esse resultado é condizente com o apontado pela literatura, especialmente no caso do Brasil (Franco et al. 2007). Nota-se que esse efeito do sexo sobre o desempenho é mais forte nas escolas privadas do que nas públicas, e que este praticamente permanece inalterado depois de controlado o efeito fixo escola, havendo uma pequena diminuição nas escolas públicas e um aumento no caso das escolas privadas. A diferença entre o coeficiente de regressão dos meninos das escolas privadas e o coeficiente de regressão dos meninos das escolas públicas é estatisticamente significativo ao nível de 0,05.

Quanto ao atraso escolar, os parâmetros estimados para as variáveis “11 anos” e “12 ou mais anos” apresentam os sinais esperados (negativos) e a significância a 1% em todas as estimações, tendo como idade de referência nove anos, que seria a de um aluno adiantado um ano na série. No caso do aluno ter 10 anos de idade, a idade correta para a série, o sinal do parâmetro é positivo e não significativo nas estimativas por Efeitos Fixos. À medida que aumenta o atraso escolar, os coeficientes estimados por EF apontam que há uma piora no rendimento dos alunos e, novamente, esse efeito negativo do atraso escolar sobre o desempenho parece ser mais severo no caso dos alunos das escolas privadas (cerca de 10 pontos a menos nas notas relativamente ao efeito sobre as notas dos alunos da escola pública). A diferença entre os coeficientes de regressão por Efeitos Fixos da variável “12 ou mais anos” entre escolas públicas

<sup>14</sup>Os resultados dos testes de Hausman são reportados nas últimas linhas das Tabelas 9 e 10.

**Tabela 9:** Proficiência em matemática 4ª série: escolas públicas e privadas (painel 1997 - 2005)

	Escolas Públicas		Escolas Privadas	
	MQO	EF	MQO	EF
Aluno				
Homem	4,989***	4,647***	6,021***	6,457***
9 anos	Base	Base	Base	Base
10 anos	1,313*	-0,341	4,003***	0,302
12 ou mais	-15,740***	-14,299***	-27,787***	-25,747***
Branco	Base	Base	Base	Base
Pardo /mulato	-0,397	2,380***	-2,631***	-0,349
Negro	-10,694***	-6,558***	-18,987***	-13,970***
Amarelo /índio	-0,535	1,424**	-1,540	-0,527
Mora pai e mãe	2,091***	0,858**	5,508***	5,026***
Escolaridade da mãe do aluno				
Nunca estudou	Base	Base	Base	Base
1à 4ª série	4,207***	2,666***	4,355	5,507*
5ª à 8ª série	6,502***	4,505***	11,785***	12,459***
Ensino médio	13,388***	9,652***	21,600***	19,528***
Faculdade	11,307***	6,109***	26,372***	20,461***
Aluno não sabe	6,008***	3,618***	17,104***	14,820***
Professor				
Homem	-3,083***	0,681	-0,846	1,161
Até 29 anos	Base	Base	Base	Base
30 a 40 anos	0,575	0,831	0,771	1,063
40 ou mais	1,369**	1,271	2,003**	-0,522
Tem ensino superior	3,754***	-0,200	3,651***	0,252
Fez curso capacitação	-0,197	0,543	1,555*	0,144
Leciona há mais de 15 anos	-0,742	-0,627	2,128***	0,129
Diretor				
Exerce função há menos de 5 anos	Base	Base	Base	Base
Exerce função entre 5 e 10 anos	-0,903	-0,589	-1,220	-1,781
Exerce função há mais de 10 anos	0,766	1,382	3,067***	-4,055**
Direção da escola há menos de 5 anos	Base	Base	Base	Base
Direção da escola entre 5 e 10 anos	1,079*	0,092	-0,198	2,795*
Direção da escola há mais de 10 anos	3,404***	0,700	-0,655	4,354**
Problema com professores faltosos	-1,081**	-0,227	-5,535***	0,331
Assumiu por concurso público, seleção, eleição.	Base	Base	Base	Base
Assumiu por indicação/ outras formas	-2,160***	0,827	-1,334**	-2,082
Projeto Pedagógico da Secretariada Educação	Base	Base	Base	Base
Projeto Pedagógico diretor e professores	1,939***	0,745	2,763***	-3,255**
Não houve Projeto Pedagógico no ano	-1,916***	0,499	-1,093	-0,278

Fonte: Cálculos próprios a partir do Censo Escolar (INEP/MEC).

Notas: Testes de diferença de médias estatisticamente significantes com níveis de significância iguais a: \*\*\* 1%, \*\* 5%, \* 10%.

**Tabela 9:** Proficiência em matemática 4ª série: escolas públicas e privadas (painel 1997 - 2005) (continuação)

	Escolas Públicas		Escolas Privadas	
	MQO	EF	MQO	EF
Escola				
Região rural	-3,582***	-	-17,725***	
Administração municipal (base: estadual)	-1,441***	-		
Biblioteca	1,882***	-1,823	-2,630***	-1,698
Laboratório de informática	1,423***	1,115	3,079***	2,115
Laboratório de ciências	7,830***	2,008	5,489***	0,714
Quadra de esportes	2,164***	0,216	3,047***	-0,465
Programa TV na escola	0,405	-1,813*	-3,565**	1,184
Merenda	4,334**	0,212	-4,678***	2,681
Professores no EF	-0,008	0,068	0,244***	0,095
Turmas na 4ª série diurnas	0,700***	-0,086	1,581***	-0,417
Turmas na 4ª série noturnas	-0,240	-0,978	1,412	-4,656
Matrículas na 4ª série	-0,030***	-0,013	-0,067***	-0,058***
D_1997	Base	Base	Base	Base
D_1999	-5,859***	-9,295**	1,024	-20,672
D_2001	-10,515***	-14,359***	5,293	-16,674
D_2003	-11,363***	-16,231***	1,874	-18,463
D_2005	-9,077***	-11,713***	6,761	-11,344
Constante do modelo	171,879***	184,689***	168,424***	223,954***
N	40.801	40.801	23.079	23.079
Grupos		1.577		971
R2 Within		0,0462		0,0575
R2 Between		0,2292		0,2081
R2 Total	0,1148	0,0802	0,1858	0,0710
Hausman	Prob > Chi2 = 0,0000		Prob > Chi2 = 0,0000	

Fonte: Cálculos próprios a partir do Censo Escolar (INEP/MEC).

Notas: Testes de diferença de médias estatisticamente significantes com níveis de significância iguais a: \*\*\* 1%, \*\* 5%, \* 10%.

**Tabela 10:** Proficiência em matemática 4ª série: escolas públicas e privadas (painel 1999-2005)

	Escolas Públicas		Escolas Privadas	
	MQO	EF	MQO	EF
Aluno				
Homem	5,388***	4,975***	6,594***	6,969***
9 anos	Base	Base	Base	Base
10 anos	1,049	-0,533	3,423***	0,057
11 anos	-9,534***	-9,043***	-5,069***	-8,952***
12 ou mais	-14,540***	-13,886***	-26,083***	-23,982***
Computador	6,124***	0,095	8,606***	1,033
Branco	Base	Base	Base	Base
Pardo /mulato	-0,631	2,169***	-2,552***	-0,214
Negro	-10,441***	-6,590***	-18,252***	-13,783***
Amarelo /índio	-0,947	0,945	-2,168*	-1,108
Mora pai e mãe	2,380***	1,051**	5,447***	5,002***
Escolaridade da mãe do aluno				
Nunca estudou	Base	Base	Base	Base
de 1ª à 4ª série	4,096***	2,668***	1,971	3,280
de 5ª à 8ª série	6,706***	4,840***	9,655***	10,010***
Ensino médio	13,446***	9,901***	19,462***	17,263***
Faculdade	10,904***	6,006***	23,754***	17,953***
Aluno não sabe	5,620***	3,429***	14,916***	12,602***
Aluno trabalha fora	-8,374***	-7,120***	-14,906***	-12,001***
Professor				
Homem	-1,920***	1,416	-1,030	0,766
Até 29 anos	Base	Base	Base	Base
30 a 40 anos	0,253	0,949	1,592**	0,555
40 ou mais	0,618	0,884	2,539**	-1,225
Tem ensino superior	3,735***	0,210	2,894***	0,540
Fez curso capacitação	-0,183	0,596	0,909	1,016
Leciona há mais de 15 anos	-0,528	-0,144	2,376***	1,123
Diretor				
Problema com professores faltosos	-1,148**	0,047	-4,970***	-0,360
Exerce função há menos de 5 anos	Base	Base	Base	Base
Exerce função entre 5 e 10 anos	-1,948***	-1,485	-1,432	-2,281
Exerce função há mais de 10 anos	1,808**	3,161	2,485**	-4,789**
Direção da escola há menos de 5 anos	Base	Base	Base	Base
Direção da escola entre 5 e 10 anos	2,376***	0,754	-1,746*	2,143
Direção da escola há mais de 10 anos	2,412**	-2,631	-0,993	4,150**
Assumiu o cargo por concurso público/seleção/eleição.	Base	Base	Base	Base
Assumiu por indicação, outras formas	-1,725***	1,019	-2,565***	-2,808*
Projeto Pedagógico da Secretaria da Educação	Base	Base	Base	Base
Projeto pedagógico diretor e professores	1,368***	1,175	3,504***	-2,854*
Não houve projeto pedagógico no ano	-1,282*	0,083	-0,459	0,119

Fonte: Estimções a partir dos dados do SAEB 1999 a 2005 e Censo Escolar 1999 a 2005.

Notas: (a) Erros padrão robustos a heterocedasticidade; (b) Erros-padrão ajustados para *clusters* de escola; (c) Não há dados para 1997. Níveis de significância dos parâmetros estimados: \* 10%; \*\* 5%, \*\*\* 1%.

**Tabela 10:** Proficiência em matemática 4ª série: escolas públicas e privadas (painel 1999-2005) (continuação)

	Escolas Públicas		Escolas Privadas	
	MQO	EF	MQO	EF
Turma				
Não teve nenhum professor de matemática no ano	Base	Base	Base	Base
Teve um professor de matemática no ano	1,684*	2,083	5,211***	2,316
Teve dois professores de matemática no ano	1,939*	0,867	2,333	2,842
Não ficou nenhum mês se professor de matemática	Base	Base	Base	Base
Ficou pelo menos 1 mês sem professor	-3,867***	-1,346	-2,434	3,423
Tamanho médio das turmas	-0,184***	-0,076	-0,174***	-0,075
Turno: menos de 4 horas aula por dia	Base	Base	Base	Base
Turno: de 4 a 5 horas aula por dia	1,609***	-0,897	4,846***	2,373
Turno: mais de 5 horas aula por dia	12,061***	0,701	11,703***	10,884**
Escola				
Região rural	-4,494***	-	-27,820***	-
Administração municipal	-0,415	-	-	-
Biblioteca	1,256***	-2,291*	-2,737**	-2,018
Laboratório de informática	-0,838	0,626	0,709	2,727
Laboratório de ciências	8,157***	3,040	3,154***	1,193
Quadra de esportes	0,759	2,344	2,780***	0,940
Internet	2,442***	-0,843	2,837***	0,149
Programa TV na escola	-0,040	-2,120**	-3,389*	1,279
Merenda	1,290	-1,866	-4,923***	-1,081
Professores no EF	-0,020	0,015	0,197***	0,099
Uso de computadores:				
Para fins pedagógicos	2,611***	0,934	3,409***	-0,594
Para uso em sala de aula	4,234**	0,459	-2,203**	1,732
Para uso na administração	3,621***	-2,180*	4,297***	-0,079
Turmas na 4ª série diurnas	-0,375	0,074	0,796*	-1,251
Turmas na 4ª série noturnas	1,032*	-1,206	6,973**	-1,103
Matriculas na 4ª série	-0,004	-0,012	-0,032**	-0,048**
d_1997	Base	Base	Base	Base
d_2001	-4,473***	-4,355***	4,748***	3,861***
d_2003	-6,705***	-6,180***	1,208	2,415
d_2005	-5,731**	-2,052	6,517***	9,188***
Constante do modelo	173,046***	179,950***	166,871***	203,868***
N	33.173	33.173	19.711	19.711
Grupos		1.413		913
R2 Within		0,0524		0,0616
R2 Between		0,2364		0,2662
R2 Total	0,1325	0,0911	0,2014	0,0888
Hausman				

Fonte: Estimacões a partir dos dados do SAEB 1999 a 2005 e Censo Escolar 1999 a 2005.

Notas: (a) Erros-padrão robustos a heterocedasticidade; (b) Erros padrão ajustados para *clusters* de escola; (c) Não há dados para 1997. Níveis de significância dos parâmetros estimados: \* 10%; \*\* 5%, \*\*\*1%.

e privadas é estatisticamente significativa a 0,05. O controle por Efeitos Fixos diminui o efeito do atraso escolar sobre o desempenho nas escolas públicas, enquanto nas escolas privadas ele aumenta no caso do efeito do aluno ter “11 anos” e diminui no caso de “12 ou mais anos”.

Quanto ao coeficiente obtido nas estimações de Efeitos Fixos relacionado à raça declarada pelo aluno, o coeficiente de “pardos ou mulatos” apresentou sinal positivo e significativo a 1%, mas apenas nas estimações feitas com escolas públicas (no painel longo e no mais curto). Além disso, nas escolas públicas esse coeficiente aumenta e muda de sinal quando se inclui o efeito escola na estimação. Para o aluno que se declara de raça “negra”, por sua vez, o coeficiente é significativo a 1% e negativo nas estimações por mínimos quadrados agrupados e Efeito Fixo tanto nas escolas públicas quanto nas privadas. Porém, vale ressaltar que nas escolas privadas ele é mais negativo (quase chegando a ser o dobro) dos coeficientes obtidos para escolas públicas. Essa diferença entre os coeficientes estimados para os alunos das escolas privadas e públicas é significativamente diferente de zero a 5%. A inclusão do Efeito Fixo atenua um pouco o impacto dessa variável nos dois tipos de escolas.

## 5.2 Resultados obtidos para o ambiente familiar e o *status* socioeconômico (SSE) dos alunos

Como mostram as estimações de Efeitos Fixos, o fato de o aluno ter computador em casa apresentou significância a 10% apenas no caso de alunos de escolas privadas no painel mais longo (1997-2005).

No que diz respeito ao ambiente familiar, a variável “mora com pai e mãe” apresentou sinal positivo e significativa a 1% e 5% nas estimações por Efeitos Fixos para escolas privadas e públicas, respectivamente. O coeficiente de Efeitos Fixos é maior (da ordem de 5,02 no painel mais longo e 5,00 no painel mais curto) e mais significativo (1%) nas escolas privadas do que os obtidos nas estimações com escolas públicas (da ordem de 0,86 no painel 1997-2005 e 1,05 no painel 1999-2005, significantes a 5%) e essa diferença é significativamente diferente de zero ao nível de confiança de 5%. A inclusão do efeito escola específico nas estimações reduz o coeficiente estimado para essa variável, tanto nas escolas públicas quanto nas particulares, em relação às estimações de MQO agrupados.

Indo ao encontro do que a literatura e outros trabalhos no mesmo gênero reportam<sup>15</sup>, os coeficientes nas estimações de Efeitos Fixos sobre o nível educacional da mãe se mostraram quase todos significantes e com o sinal positivo. Porém, há algumas diferenças nos resultados obtidos com escolas públicas e com escolas privadas. A escolaridade da mãe contribui para um melhor desempenho dos alunos nas provas. Nas escolas privadas o fato de a mãe ter feito apenas o ensino primário (de 1ª à 4ª série) só apresentou significância no painel mais longo a 10%, enquanto nas escolas públicas os resultados obtidos para essa variável são todos significativos a 1%. Além disso, pela magnitude dos coeficientes nas estimações de Efeitos Fixos, os alunos de escolas privadas se beneficiam mais do nível educacional da mãe do que os alunos das escolas públicas, enquanto se observa apenas uma pequena redução na importância desses coeficientes quando se inclui na regressão o efeito fixo de escola. A relação condicional positiva entre o estudo da mãe e o desempenho do aluno

<sup>15</sup>Ver, por exemplo, Menezes-Filho (2007), Felício & Fernandes (2006).

crece à medida que o nível educacional da mãe aumenta, principalmente no caso das escolas privadas.

Por fim, no painel mais curto, tanto nas estimações para escolas privadas quanto para as públicas, o coeficiente da variável “aluno trabalha fora” apresentou sinal negativo e altamente significativo (1%) nas estimações por MQO agrupados e Efeitos Fixos. Os alunos das escolas particulares parecem sofrer mais prejuízo em seu aprendizado quando trabalham fora do que os alunos das escolas públicas. O coeficiente de “aluno trabalha fora” é da ordem de  $-12$  nas escolas privadas, e de  $-7,12$  nas escolas públicas, e essa diferença é significativamente diferente de zero ao nível de confiança de 5%.

### **5.3 Resultados obtidos para as características dos professores**

Quanto aos efeitos de algumas características observáveis dos professores, seus coeficientes se mostraram todos não significativos quando é incluído o efeito escola-específico na regressão. Principalmente no que diz respeito ao nível educacional do professor, os resultados só se mostraram altamente significativos e positivos nas estimações por MQO. De acordo com o que defende Hanushek (1995), o nível educacional do professor e demais características observáveis, como experiência e capacitação, não captam a qualidade do professor, isto é, não espelham se o professor é capaz de agregar maior aprendizado aos seus alunos. Nesse caso, os coeficientes obtidos por MQO agrupados para essas variáveis estariam refletindo apenas relações espúrias. Porém, é possível argumentar que o nível educacional do professor exerce influência positiva sobre o aprendizado dos alunos, mas que o método de Efeitos Fixos inclui no termo escola-específico não observado muita informação, inclusive sobre o efeito fixo não observado da qualidade do professor, que pode estar correlacionado com a sua educação e experiência, mas que inclui outras características não observáveis. Por exemplo, as escolas têm processos de escolha dos professores que tomam por base o nível educacional dos mesmos, bem como uma série de outras informações como habilidade, histórico profissional, relações de amizade ou “quem indique” e, portanto, tudo isso pode estar sendo captado pelo efeito específico da escola nas regressões por Efeitos Fixos.

### **5.4 Resultados obtidos para as características dos diretores**

Nas escolas privadas, a única característica observável dos diretores que apresentou sinal negativo e significância de 5% nas estimações por Efeitos Fixos foi o fato do diretor “exercer a função há mais de 10 anos”. Além disso, nas escolas privadas em que “o projeto pedagógico foi desenvolvido pelo diretor e professores”, e não pela Secretaria de Educação, apresentou um coeficiente de Efeitos Fixos com sinal negativo e significativo a 5%, no painel mais longo, e a 10% no painel mais curto. A variável “assumiu o cargo por indicação ou outras formas” apresentou parâmetro com sinal negativo e significativo a 10%, mas apenas na estimação com o painel mais curto. No que diz respeito às características dos diretores de escolas públicas, nenhuma delas se mostrou significativa nas estimações de Efeitos Fixos.

### **5.5 Resultados obtidos para as características das turmas**

No painel mais curto, de 1999 a 2005, foi possível incluir variáveis com características das turmas dos alunos que fizeram a prova do SAEB. Tanto nas

estimações com as escolas privadas quanto nas com escolas públicas, as estimasções com Efeitos Fixos, em geral, não deram resultado significativo. Nas escolas privadas, o coeficiente da variável que indica que as turmas têm mais de 5 horas-aula por dia deu positivo e significativo a 1% tanto na estimasção por MQO agrupados (coeficiente igual a 11,7) quanto na estimasção por Efeitos Fixos (coeficiente igual a 10,88). Nenhum outro resultado por Efeitos Fixos deu significativo no caso das escolas privadas.

### 5.6 Resultados obtidos para as características das escolas

No que diz respeito aos efeitos das características das escolas, nas escolas privadas o único efeito que se mostrou significativo nos dois modelos estimados no painel mais longo foi o de “número de matrículas na 4ª série”, que apresentou sinal negativo e significância a 1% nas estimasções MQO agrupados e Efeitos Fixos (no painel mais curto as mesmas estimasções se mostraram significantes a 5%). Este pode ser um indício de que há um efeito escala negativo a partir de determinado ponto nas escolas privadas, em que os recursos por alunos matriculados na série em questão se tornam mais escassos. Vale notar, porém, que a magnitude desses coeficientes é bem pequena, ficando entre  $-0,067$  e  $-0,032$  apenas.

Nas estimasções com escolas públicas e dados do painel mais curto (1999-2005), o fato da escola se situar na região rural apresentou um sinal negativo na estimasção por MQO agrupados, resultado já encontrado na literatura nacional sobre o tema. Nas estimasções por Efeitos Fixos, os efeitos de a escola ter “biblioteca”, participar do programa do MEC “TV na Escola” e ter computadores para “uso na administração” apresentaram sinais negativos e significância a 10%, 5% e 10%, respectivamente. Isso pode decorrer do problema de “causalidade reversa” nas variáveis. Por exemplo, exatamente nas escolas que apresentavam um desempenho ruim, o governo realizou investimentos em bibliotecas, compra de computadores e inclusão da escola em programas educacionais para tentar reverter o mau desempenho. É possível que se tenha captado nas estimasções exatamente um ponto na curva da qualidade da escola decrescente, quando tais políticas de insumo foram adotadas.

## 6 Comentários Finais

O que se pôde observar nas análises realizadas é que apenas algumas variáveis se mostram altamente significativas para explicar o desempenho escolar, independentemente do ano e da extensão do painel utilizado, e que essas variáveis se resumem praticamente às características dos alunos, o seu atraso escolar, e ao status socioeconômico de suas famílias. Esperava-se que, por se trabalhar com um painel de dados com um número maior de anos, fosse possível captar melhor a influência que características observáveis das escolas, diretores, turmas e professores exercem sobre a proficiência dos alunos, principalmente por meio do modelo Efeitos Fixos que controla as características não observáveis específicas de cada escola e constantes no tempo. Isso poderia ser muito importante em termos de recomendações de políticas para melhorar a qualidade do ensino no país bem como a qualidade dos gastos públicos na educação. Essa expectativa, porém, não se comprovou nas análises acima realizadas.

Um resultado importante verificado nas análises é que em geral os efeitos das variáveis que entraram nas regressões, tanto positivos quanto negativos,

apresentaram uma importância maior no desempenho escolar dos alunos das escolas particulares do que no desempenho dos alunos das redes públicas de ensino (escolas estaduais e municipais). O atraso escolar é um exemplo nesse sentido, pois as estimações indicam que seus efeitos são mais severos sobre o desempenho dos alunos das escolas privadas do que sobre os das escolas públicas, com uma diferença entre os coeficientes de regressão estatisticamente significativa a 0,05 em se tratando dos alunos atrasados dois anos na série (12 anos). Seria isso um indicador de melhor ou pior qualidade nas escolas particulares? Talvez as escolas particulares sejam menos severas do que as escolas públicas na hora de avaliar se o aluno adquiriu os conhecimentos necessários para evoluir de série, deixando para trás apenas aqueles que se saem realmente muito mal. Porém, dado que o desempenho das escolas privadas é, em média, bem melhor do que das escolas públicas (ver Tabela 5), isso pode também indicar que nas escolas privadas os alunos conseguem aprender mais do que nas escolas públicas, ficando para trás alunos que, se avaliados em uma escola pública, tenderiam a ser aprovados.

Dois resultados que podem espelhar práticas educacionais importantes emergem dos resultados encontrados no caso das escolas privadas. A primeira delas é de que um turno superior a 5 horas-aula pode aumentar o aprendizado dos alunos. A segunda seria a de garantir uma taxa adequada de insumos por aluno, uma vez que parece haver um efeito escala negativo quando aumenta muito o número de matrículas na quarta série. Por que os efeitos dessas duas variáveis não se mostraram significativos para as escolas públicas? Talvez variáveis não observáveis nas escolas públicas podem estar impedindo que essas políticas apresentem resultados positivos<sup>16</sup>.

A maioria das variáveis nessa análise se mostrou determinante para o aprendizado nas escolas públicas, porém, não estão diretamente associadas a políticas de educação. Essa constatação vai ao encontro do que Hanushek (Op.cit) e que parte da literatura internacional sobre o tema defendem: que as características que fazem uma boa escola, um bom professor ou um bom diretor não são passíveis de mensuração e não se explicam por características observáveis, tais como os anos de escolaridade, a idade ou seus anos de experiência.

Por outro lado, talvez o método de estimação não seja o adequado para se medir o impacto das variáveis regressoras de interesse, como defendido por Lee (2001). Será que o uso de modelos hierárquicos traria resultados diferentes? Por enquanto, a única conclusão que os resultados nos permitem inferir é que as condições socioeconômicas dos alunos e suas famílias são as grandes responsáveis pelo desempenho dos alunos no caso aqui estudado (alunos da 4ª série avaliados nas provas de matemática), e que a estratificação social que existe no país replica-se no sistema educacional. Este, por sua vez, não consegue quebrar as barreiras da estratificação, tornando-se mais uma força no sentido da reprodução da desigualdade.

---

<sup>16</sup>Para uma análise mais profunda sobre o custo-efetividade dessas políticas, ver Oliveira (2008), que encontra efeitos significativos sobre o desempenho dos alunos da 4ª série em matemática de escolas públicas da jornada escolar e do tamanho da escola, utilizando a técnica de estimador de *matching* e dados do SAEB 2005.

## Referências Bibliográficas

- Alves, M. T. G. & Soares, J. F. (2007), 'As pesquisas sobre o efeito das escolas: contribuições metodológicas para a sociologia da educação', *Sociedade e Estado* 22(2), 435–473.
- Angrist, J. D. & Krueger, A. B. (1999), Empirical strategies in labor economics, in O. Arshenfelter & D. Card, eds, 'Handbook of labor economics', Vol. 3, Elsevier.
- Biondi, R. L. & Felício, F. (2007), *Atributos escolares e o desempenho dos estudantes: uma análise de painel dos dados do SAEB*, 1 edn, INEP, Brasília.
- Felício, F. & Fernandes, R. (2006), 'O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de são paulo', Dissertação.
- Franco, C. et al. (2007), 'Qualidade e equidade em educação: reconsiderando o significado de fatores intra-escolares', *Avaliação de Políticas Públicas Educacionais* 15(55), 277–298.
- Glewwe, P. & Kremer, M. (2006), Schools, teachers, and education outcomes in developing countries, in E. Hanushek & F. Welch, eds, 'Handbook of the Economics of Education', Vol. 2, Elsevier.
- Hanushek, E. (1995), 'Interpreting recent research on schooling in developing countries', *World Bank Research Observer* 10(2), 227–246.
- Hanushek, E. (2002), Teacher quality, in L. E. W. M. IZUMI, ed., 'Teacher Quality', Hoover Institution Press.
- Hanushek, E. (2005a), 'Teachers, school and academic achievement', *Econometrica* 73(2), 417–458.
- Hanushek, E. (2005b), 'Why quality matters in education', *Finance and development* 42(2), 15–19.
- Hanushek, E. & Pace, R. R. (1995), 'Who chooses to teach (and why)?', *Economics of education review* 14(2), 101–117.
- INEP (2009a), 'Censo escolar', <http://portal.inep.gov.br/microdados>.
- INEP (2009b), 'Sistema de avaliação da educação básica - saeb', <http://portal.inep.gov.br/microdados>.
- Kremer, M. (1995), 'Randomized evaluations of educational programs in developing countries: some lessons', *American Economic Review* 93(2), 102–115.
- Lazear, E. (2003), 'Teacher incentives', *Swedish Economic Policy Review* 10(3), 179–214.
- Lee, V. E. (2001), 'Using multilevel methods to investigate research questions that involves nested data: examples from education', *Estudos em avaliação educacional* (24), 47–68.

Menezes-Filho, N. A. (2007), 'Os determinantes do desempenho escolar no Brasil.', <https://www.todospelaeducacao.org.br//arquivos/biblioteca/f4e8070a-8390-479c-a532-803bbf14993a.pdf>.

Mulligan, C. B. (1999), 'Galton versus the human capital approach do inheritance', *Journal of Political Economy* 107(6), 184–224.

Murnane, R. J. et al. (2000), 'How important are cognitive skills of teenagers in predicting subsequent earnings?', *Journal of Policy Analysis and Management* 19(4), 547–568.

Oliveira, J. M. (2008), 'Custo efetividade de políticas de redução do tamanho da classe e ampliação da jornada escolar: uma aplicação de estimadores de matching', Dissertação.