

Diferenças intertemporais na média e distribuição do desempenho escolar no Brasil: o papel do nível socioeconômico, 1997 a 2005

Clarissa Guimarães Rodrigues*
Eduardo Luiz Gonçalves Rios-Neto**
Cristine Campos de Xavier Pinto***

É consenso na literatura de avaliação educacional que o nível socioeconômico dos alunos é o fator mais importante para explicar, em determinado ponto no tempo, as variações nos resultados escolares dos alunos avaliados pelo Sistema de Avaliação da Educação Básica – Saeb. Não se sabe, porém, em que medida a mudança temporal nas condições socioeconômicas influenciou a mudança na média e na distribuição do desempenho escolar entre os diversos ciclos do Saeb. Para investigar esta questão, utilizamos o método de decomposição contrafactual desenvolvido por Juhn, Murphy e Pierce (1993), que permite isolar a contribuição da variação na composição e retorno do nível socioeconômico sobre a variação na média e distribuição do desempenho escolar. Foram empregados os dados do Saeb de 1997 a 2005, para a coorte de alunos da 4ª série do ensino fundamental avaliados em matemática. Os principais resultados mostram que o aumento na cobertura escolar no final dos anos 1990 reduziu o nível socioeconômico médio dos alunos e colaborou para diminuir a média e elevar a desigualdade entre os resultados escolares (efeito composição). Ao mesmo tempo, houve redução na sensibilidade do desempenho escolar ao nível socioeconômico (efeito retorno). Este achado tem implicações ambíguas: por um lado, contribui para reduzir o aproveitamento escolar médio e, por outro, colabora para equalizar a distribuição dos escores entre 1997 e 2005.

Palavras-chave: Desempenho escolar. Nível socioeconômico. Decomposição contrafactual.

Introdução

Estudos quantitativos com foco na avaliação da educação escolar acumulam, há muitos anos, evidências que comprovam a importância das origens sociais no processo de construção do sucesso escolar de crianças e jovens. Nos Estados Unidos, pesquisas sobre avaliação da educação escolar têm sido realizadas desde o início do século XX

(TEDDLIE; STRINGFIELD, 2007; MADAUS et al., 2008). Entretanto, a consolidação de uma linha de pesquisa quantitativa focada na avaliação da relação entre os atributos escolares e o desenvolvimento das habilidades cognitivas dos estudantes ocorreu apenas após a publicação do conhecido Relatório Coleman, em 1966 (BOGOTCH et al., 2007; TEDDLIE; STRINGFIELD, 2007; CREEMERS, 2007; BROOKE; SOARES, 2008).

* Doutora em Demografia e pesquisadora do Cedeplar/UFMG.

** Professor titular do Departamento de Demografia e pesquisador do Cedeplar/UFMG.

*** Professora do Departamento de Economia da EESP/FGV.

No Brasil, os estudos direcionados para avaliação do desempenho escolar iniciaram-se após a consolidação do Sistema de Avaliação da Educação Básica – Saeb, no início da década de 1990. O Saeb consiste em uma avaliação, em larga escala, da qualidade do ensino oferecido pelo sistema educacional brasileiro, com base na aferição do desempenho dos alunos nos exames de proficiência para uma determinada etapa da educação formal. É formado também por questionários que coletam informações contextuais dos estudantes e estabelecimentos de ensino (incluindo dados sobre turmas, professores, diretores, instalações físicas e práticas internas das escolas).

Desde 1995, as informações sobre o desempenho dos alunos nos exames de proficiência produzidos pelo Saeb passaram a ser comparáveis entre os anos e entre as séries escolares. Esta comparação foi possível devido à mudança metodológica introduzida na construção dos instrumentos e atribuição dos escores que medem o aprendizado do aluno, dada pela substituição da Teoria Clássica dos Testes (TCT) pela Teoria de Resposta ao Item (TRI). Com a disponibilização e a composição de uma série histórica bienal, os dados produzidos pelo Saeb possibilitaram o monitoramento do nível e da tendência do desempenho escolar médio dos estudantes brasileiros das 4^a e 8^a séries do ensino fundamental e 3^a série do ensino médio.

A constatação de uma tendência declinante dos níveis de desempenho alcançado pelos estudantes da educação básica, principalmente após a divulgação dos dados de 1999, desencadeou uma onda de reflexões e estudos acerca do problema da baixa qualidade do ensino no Brasil (BRASIL, 2000; NERI; CARVALHO, 2002; FERNANDES; NATENZON, 2003; SOUZA, 2006; ALVES, 2007). Uma das explicações para este fenômeno se pautou no aumento da heterogeneidade do público escolar, em virtude das políticas de ampliação do acesso ao ensino e de correção do fluxo escolar, implementadas na segunda metade da década de 1990. Entretanto, os

resultados dos estudos não são conclusivos e apontam para a necessidade de novas investigações.

O presente trabalho pretende contribuir para esta literatura, procurando investigar o efeito das mudanças ocorridas nas características individuais, familiares e escolares dos alunos da 4^a série do ensino fundamental sobre as alterações ocorridas no desempenho escolar entre 1997 e 2005. Utilizando como referência estudos já consolidados na literatura sobre este tema,¹ pode-se afirmar que o aproveitamento escolar de um estudante é produzido por diferentes insumos, que incluem as características observáveis e não-observáveis do aluno e da escola por ele frequentada. Cada um desses insumos possui determinada produtividade ou taxa de retorno, que é necessária para que o mesmo se transforme em proficiência escolar. Esta relação de produtividade entre os insumos individuais, familiares e escolares e o produto final (no caso, o desempenho escolar) constitui a essência da função de produção educacional.

Tomando como base a função de produção educacional, consideramos que as variações na proficiência escolar observadas entre 1997 e 2005 podem ser explicadas por três fontes: mudanças na composição das características (Xs) observáveis relacionadas ao aluno e à escola – efeito composição; alterações na produtividade ou taxa de retorno (β s) dessas características – efeito retorno; e mudanças na composição e retorno das características não observadas dos alunos e das escolas – efeito resíduo.

Para identificar estes efeitos, utilizamos o método de decomposição baseado em microssimulação contrafactual desenvolvido por Juhn, Murphy e Pierce (1993), bastante empregado nos estudos sobre as desigualdades de renda.² Com esta metodologia, foi possível identificar e medir os componentes que explicam a diferença na média e na distribuição (dada pela diferença entre os percentis: 90-10, 90-50, 50-10) do desempenho escolar, no período de 1997 a 2005.

¹ Ver Fletcher (1998); Soares, César e Mambrini (2001); Ferrão e Fernandes (2001); Albernaz et al. (2002); Soares (2005); para citar alguns.

² Como exemplo de trabalhos recentes que utilizam esta metodologia, ver Soares (2004), Azevedo e Foguel (2006) e Ulyssea (2007).

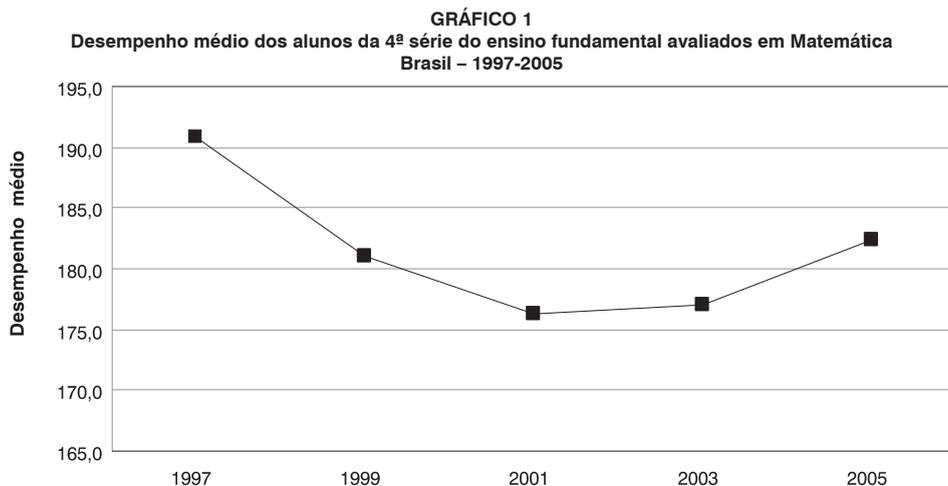
A ideia em transpor os métodos desenvolvidos na área de economia do trabalho para os estudos educacionais justifica-se pela similaridade existente na natureza das variáveis-resposta utilizadas em ambas as áreas: salário e proficiência. Estatisticamente, essas variáveis são classificadas como quantitativas e contínuas e podem ser consideradas importantes preditoras do bem-estar individual. Entretanto, diferentemente dos estudos econômicos – que tradicionalmente consideram apenas as variáveis individuais (por exemplo, escolaridade e experiência dos trabalhadores) como fatores associados ao rendimento do trabalho –, nos estudos educacionais tanto as variáveis desagregadas por indivíduo quanto aquelas agregadas por escola são importantes preditoras do desempenho escolar.

Para lidar com a hierarquia presente nos dados educacionais e, ao mesmo tempo, decompor a variação do desempenho nas unidades micro (aluno) e macro (escola), adotamos a estratégia de implementar as regressões separadamente em cada uma das unidades de análise: aluno e escola. Embora reconheçamos que esta estratégia metodológica não seja a melhor opção em estudos que envolvem dados de natureza hierárquica, seu uso justifica-se pela possibilidade de identificar cada um dos efeitos – composição, retorno e

resíduo – em ambos os níveis. Tendo em vista que nos exercícios de decomposição estamos interessados apenas na magnitude dos coeficientes e não em sua variância, esta alternativa se encaixa ao propósito do presente estudo, uma vez que os coeficientes gerados pelo método dos mínimos quadrados ordinários e pelo modelo hierárquico linear são muito similares³ (RAUDENBUSH; BRYK, 2002; BICKEL, 2007). Portanto, apesar das limitações impostas pelo uso dos modelos de regressão linear clássica aos dados educacionais, consideramos que este estudo contribui de forma relevante para a literatura de avaliação educacional, pois ainda não há trabalhos que procuram identificar e medir os efeitos descritos anteriormente.

Desempenho escolar e nível socioeconômico: como estas medidas variaram entre 1997 e 2005?

Desde o início da avaliação do desempenho escolar dos estudantes da educação básica brasileira, na década de 1990, nota-se uma preocupação com os baixos níveis alcançados pelos alunos, principalmente a partir da constatação de uma tendência declinante deste indicador da qualidade da educação entre 1997 e 2005, apontada pelos dados do Saeb, como mostra o Gráfico 1.



Fonte: Inep. Sistema de Avaliação da Educação Básica –Saeb, 1997, 1999, 2001, 2003 e 2005.

³ Em Rodrigues (2009), há uma descrição detalhada sobre as diferenças e similaridades existentes na estimação dos parâmetros via MQO e HLM. A autora estimou vários modelos utilizando ambos os métodos com o intuito de comparar os coeficientes. Os resultados mostram que os coeficientes são praticamente iguais.

Esta evidência suscita questões acerca dos fatores que podem ter influenciado os resultados educacionais dos estudantes neste período. A primeira justificativa dada para a piora no aproveitamento médio dos alunos esteve associada ao aumento da cobertura do ensino fundamental, que ocorreu, principalmente, devido à inclusão de crianças com baixo nível socioeconômico⁴ no sistema de ensino (SOUZA, 2006).

Embora o discurso oficial seja desprovido de evidências empíricas mais aprofundadas e se respalde apenas nos achados internacionais, o argumento utilizado se sustenta nos resultados encontrados pela literatura sobre os fatores associados ao desempenho escolar no Brasil (FLETCHER, 1998; CÉSAR, SOARES; MAMBRINI, 2001; FERRÃO; FERNANDES, 2001; ALBERNAZ et al., 2001; SOARES, 2005).

Há consenso de que o nível socioeconômico dos alunos é o fator mais associado ao desempenho escolar. Em outras palavras, pais com elevado (baixo) capital econômico e cultural são mais (menos) propensos a incentivarem e valorizarem o estudo de seus filhos. Assim, o aumento na proporção de crianças com baixo nível socioeconômico no sistema de ensino levaria ao aumento na proporção de piores resultados escolares e, consequentemente, reduziriam o desempenho médio global.

A mudança na composição social dos estudantes é apenas uma das faces das alterações ocorridas no perfil do alunado. Estudos recentes têm enfatizado a melhoria do fluxo escolar, verificada na década de 1990, como um dos fatores desencadeadores de mudanças na composição etária do grupo de alunos que frequentam determinada série escolar (NERI; CARVALHO, 2002; FERNANDES; NATENZON, 2003; ALVES, 2007).

Com o objetivo de evitar os vieses de seletividade na medição da qualidade

do ensino produzidos pela expansão das matrículas e, principalmente, pelo aumento nas taxas de aprovação, estes estudos procuram comparar o desempenho escolar entre coortes, em vez de compará-lo entre as séries avaliadas pelo Saeb. A ideia é que, ao comparar o desempenho entre gerações sucessivas – por exemplo, entre as coortes de crianças com dez anos de idade ao longo dos anos em que o Saeb foi aplicado –, obtêm-se estimativas das habilidades e competências cognitivas livres do efeito de mudanças na composição etária das turmas que frequentam a 4ª série do ensino fundamental.

Dado que uma parcela da população de crianças com dez anos não foi avaliada pelo Saeb, seja por não frequentar a escola, seja por estar atrasada/adiantada em relação à série escolar incluída na avaliação, os pesquisadores desenvolveram metodologias que combinam os dados do Saeb com os de outras pesquisas⁵ para obter as informações faltantes sobre o desempenho escolar da coorte em análise.

Neri e Carvalho (2002) constataram aumento de 3,48 pontos, em Português, e 3,67 pontos, em Matemática, no desempenho escolar médio das coortes de dez anos de idade, entre 1995 e 2001. Já Fernandes e Natenzon (2003) verificaram crescimento no desempenho escolar do 7º decil das coortes de dez anos, entre 1995 e 1999: 13,16 pontos, em Português; e 18,34 pontos, em Matemática. Por fim, Alves (2007) observou redução de 7,5 pontos no desempenho escolar médio em Matemática, para as estas mesmas coortes, entre 1999 e 2003.

Como é possível notar, os resultados encontrados nos três trabalhos citados divergem em termos de magnitude e direção, no que tange às variações no desempenho escolar ao longo dos anos. Enquanto Neri e Carvalho (2002) e Fernandes e Natenzon

⁴ Souza (2005) mostra que a taxa de atendimento das crianças de 7 a 14 anos residentes em domicílios cuja renda situa-se no 1º quinto da distribuição (20% mais pobres) aumentou de 75% para 94%, entre 1992 e 2000. Neste mesmo período, a taxa de atendimento das crianças de 7 a 14 anos das famílias que compõem o estrato dos 20% mais ricos cresceu apenas 2 pontos percentuais, passando de 97% para 99%.

⁵ Neri e Carvalho (2002) e Fernandes e Natenzon (2003) utilizam os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) nos anos em que o Saeb foi aplicado. Além da PNAD, Alves (2007) considera também a base de dados do Nova Escola de 2004.

(2003) concluem que, caso não houvesse o efeito de mudanças na composição etária, o desempenho escolar teria aumentado no período em análise (um crescimento modesto, no primeiro trabalho, e outro expressivo, no segundo), Alves (2007) revela uma redução neste indicador de qualidade do ensino.

As diferenças nos resultados encontrados pela literatura, no que se refere ao efeito da heterogeneidade do público escolar sobre as variações nos resultados dos exames de proficiência, apontam para a necessidade de compreender melhor os fatores que estiveram por trás das variações no desempenho escolar, principalmente em um contexto marcado por reformas educacionais e mudanças no perfil da população que frequenta o sistema de ensino.

O Gráfico 2 mostra que, entre 1997 e 1999, período marcado pelo expressivo declínio do desempenho escolar médio, houve um aumento de 4,59% nas matrículas totais na 4ª série do ensino fundamental. Pode-se imaginar o crescimento das matrículas sendo alimentado por dois efeitos: demográfico e de políticas educacionais.

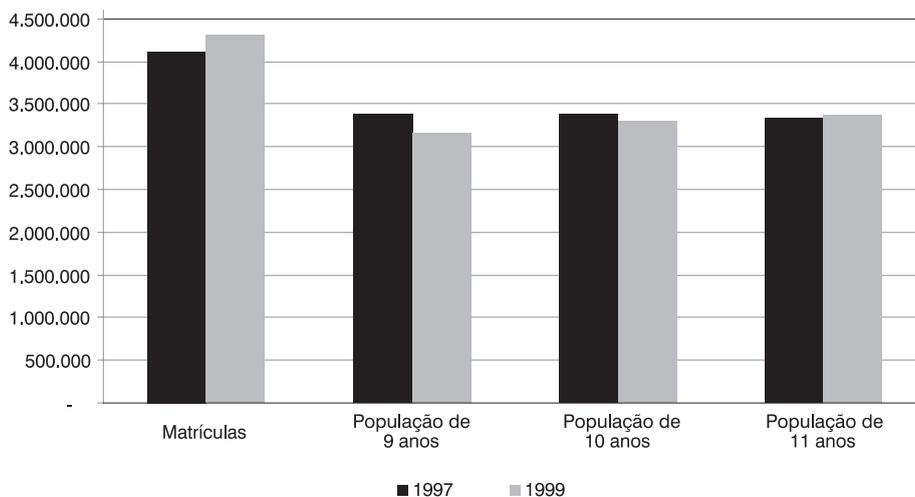
O efeito demográfico seria dado pelo aumento no tamanho das coortes em idade de frequentar este nível de ensino. Uma

aproximação deste efeito pode ser vista pela variação no tamanho das coortes de 9, 10 e 11 anos, uma vez que aproximadamente 80% dos alunos matriculados na 4ª série do ensino fundamental encontram-se nesta faixa etária.

Entretanto, o Gráfico 2 revela uma redução no tamanho das coortes de 9 e 10 anos e um pequeno aumento na de 11 anos. Dessa forma, o efeito demográfico parece não explicar a grande expansão das matrículas ocorrida entre 1997 e 1999.

O crescimento das matrículas nesta série escolar parece ter sido alimentado pelas políticas educacionais implementadas no período. Conclusão similar é encontrada no estudo de Soares e Souza (2003). Podemos citar, por exemplo, a implementação do Fundo de Manutenção e Desenvolvimento da Educação Fundamental e Valorização do Magistério (Fundef), em 1998; a criação do Programa Nacional de Garantia de Renda Mínima, iniciado em 1997, que tinha como um dos objetivos assegurar o acesso e a permanência de crianças e adolescentes de 7 a 14 anos de menor renda no sistema de ensino; e a criação da Lei de Diretrizes e Bases da Educação (LDB), que estabeleceu, entre outras diretrizes, a possibilidade de as escolas do ensino fundamental adotarem o

GRÁFICO 2
Matrículas na 4ª série do ensino fundamental e população de 9 a 11 anos de idade
Brasil – 1997-1999



Fonte: Inep. Censo Escolar; IBGE. Pesquisa Nacional por Amostra de Domicílios, 1997 e 1999.

sistema de ciclos e progressão continuada e, conseqüentemente, reduzirem as taxas de abandono e evasão escolares.

Nesse sentido, é coerente imaginar que a inclusão tenha ocorrido entre as camadas mais pobres da população. De fato, no período posterior às reformas educacionais implementadas na década de 1990, houve um aumento na proporção de alunos com mães menos escolarizadas, o que sinaliza uma mudança na composição socioeconômica dos estudantes. Observa-se, na Tabela 1, que a proporção de alunos na 4ª série do ensino fundamental cujas mães nunca estudaram aumentou entre 1997 e 2001, enquanto a daqueles com mães mais escolarizadas se reduziu. Este resultado é mais evidente entre os estudantes situados no 1º decil da distribuição da proficiência.

A revisão da literatura e as estatísticas apresentadas nesta seção mostram que a composição do público escolar brasileiro modificou-se ao longo dos últimos anos, em decorrência dos efeitos de políticas educacionais de ampliação do acesso escolar e manutenção das crianças e jovens no sistema de ensino. A inclusão de alunos com menor *background* familiar e a manutenção daqueles com maiores dificuldades de aprendizado estão entre os fatores que afetam diretamente o resultado educacional global. Como o Saeb coleta informações de período e não de coorte, mudanças na

heterogeneidade dos atributos individuais, familiares e escolares que se associam ao desempenho dos alunos, sejam estas observáveis ou não, podem afetar de forma sistemática as medidas dos resultados educacionais em cada período. Surge, portanto, a necessidade de estudar alguns aspectos pouco explorados na literatura, como o efeito da dinâmica das características dos alunos e das escolas sobre a variação no desempenho escolar nos últimos anos.

Dados

Base de dados

Os dados empregados neste estudo foram coletados pelo Sistema de Avaliação da Educação Básica (Saeb) e fornecidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Utilizaram-se as informações obtidas nos ciclos de avaliação realizados em 1997, 1999, 2001, 2003 e 2005, para a 4ª série do ensino fundamental no Brasil.⁶

As bases de dados do Saeb contemplam dois tipos de informações: proficiência dos estudantes em Língua Portuguesa e Matemática; e características contextuais dos alunos e das escolas por eles frequentadas (incluindo as informações sobre seus professores e diretores).

TABELA 1
Proporção de alunos da 4ª série do ensino fundamental avaliados em Matemática, segundo escolaridade da mãe Brasil – 1997-2005

Escolaridade da mãe	Em porcentagem				
	1997	1999	2001	2003	2005
Total de alunos					
Mãe que nunca estudou	6,27	7,16	8,57	5,14	3,62
Mãe com nove anos ou mais de estudo	21,27	16,58	15,34	24,6	23,54
Alunos do 1º decil					
Mãe que nunca estudou	10,44	13,06	16,29	10,61	8,15
Mãe com nove anos ou mais de estudo	13,35	7,86	9,11	13,43	10,56
Alunos do 10º decil					
Mãe que nunca estudou	0,87	1,53	1,24	0,42	0,45
Mãe com nove anos ou mais de estudo	46,31	42,27	47,33	59,27	53,33

Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

⁶ A não inclusão do ano de 1995 na análise deve-se à inexistência de informações referentes à escolaridade dos pais dos alunos avaliados na 4ª série do ensino fundamental.

A amostra utilizada neste trabalho, apresentada na Tabela 2, abrange os alunos da 4ª série do ensino fundamental matriculados nas escolas públicas (exceto as federais) e particulares localizadas na área urbana.⁷

Como é possível notar, a partir de 1999 houve aumento relevante no número de escolas e redução no número máximo de alunos por escola. Tal fato se explica pelas modificações ocorridas no plano amostral do Saeb. Conforme descrito por Franco (2001), a partir de 1999, o processo de amostragem foi modificado. Em vez de a amostra ser coletada por conglomerado em dois estágios (sorteio de municípios no primeiro estágio e de escolas no segundo), as escolas passaram a ser sorteadas diretamente com o objetivo de se obter maior espalhamento da amostra.

Contudo, esta mudança metodológica não interfere na comparabilidade dos resultados quando são utilizados os pesos amostrais, pois, segundo o relatório técnico da amostra do Inep, presente no Dicionário de Dados do Saeb de 2001, “o seu uso garante a compensação dos efeitos na amostragem com taxas diferenciadas nos diversos estratos da pesquisa”.

Optamos por trabalhar com a amostra de alunos avaliados em Matemática, devido

ao caráter mais tipicamente escolar do conhecimento desta disciplina.

A escolha em trabalhar com os dados da 4ª série do ensino fundamental justifica-se por ser esta a série escolar, entre as três avaliadas pelo Saeb, que apresenta menor seletividade, uma vez que a cobertura está praticamente universalizada, além de ter sido a mais afetada pelo processo de expansão do ensino, quando comparada à 8ª série do ensino fundamental e à 3ª do ensino médio. Outra questão está relacionada ao fato de que, como os alunos ainda estão no início da trajetória escolar, o controle pelo conhecimento prévio do estudante é maior, o que minimiza os problemas derivados do uso de informações sobre o desempenho do aluno em um único período, quando não se dispõem de informações sobre o valor adicionado.

Variáveis selecionadas e estatísticas descritivas

As variáveis relacionadas ao aluno e à escola foram selecionadas com base no referencial teórico proposto por Franco et al. (2003), na literatura especializada e nas possibilidades presentes nos questionários contextuais do Saeb.

TABELA 2
Amostra de escolas e alunos da 4ª série do ensino fundamental avaliados em Matemática e estatística descritivas dos alunos por escola
Brasil – 1997-2005

Anos	Amostra de alunos	Amostra de escolas	Alunos por escola			
			Média	Desvio-padrão	Mínimo	Máximo
1997	18.588	698	26,40	13,46	2	80
1999	16.811	2.898	5,70	3,49	1	20
2001	50.782	3.551	14,21	8,83	1	52
2003	40.596	2.915	13,92	8,63	1	55
2005	37.719	2.508	14,90	8,37	1	52

Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

⁷ A exclusão dos alunos de escolas rurais e federais surgiu da necessidade de compatibilizar a amostra ao longo dos ciclos de avaliação do Saeb, para garantir a comparabilidade dos resultados. Em relação às escolas rurais, em 1999 e 2001, foram pesquisadas apenas aquelas localizadas nos Estados do Nordeste, em Minas Gerais e no Mato Grosso do Sul. Quanto às escolas federais, nos ciclos do Saeb de 1997, 1999 e 2001, as mesmas não entraram na amostra (BRASIL, 2006). Alguns dados da participação das escolas rurais no total de alunos: 1997 (2,26%); 1999 (5,29%); 2001, 2003 e 2005 (0,00%). No que tange à participação das escolas federais no total de alunos, em 2003 e 2005, este número foi próximo a 0%.

Desde a implementação do Saeb, os questionários contextuais têm passado por modificações que visam o aprimoramento das medidas dos fatores individuais, familiares e escolares que se associam ao desempenho escolar dos alunos. Por esse motivo e, também, pelo fato de compararmos modelos iguais entre os anos, procuramos selecionar as variáveis que se repetem nos questionários dos ciclos do Saeb de 1997, 1999, 2001, 2003 e 2005 e que apresentam um padrão nos itens de resposta.

Para a construção dos modelos, nos respaldamos nos resultados da literatura e selecionamos as variáveis comumente utilizadas no nível do aluno e da escola, seguindo a opção de construir modelos parcimoniosos.

Nível socioeconômico

Devido à importância desta variável na análise que será feita a seguir, optamos por descrever, de forma sucinta, a estratégia adotada para construir o nível socioeconômico dos estudantes.

A condição socioeconômica dos alunos não é uma variável diretamente observada e, portanto, não está disponível nos dados produzidos pelo Saeb. Dada a importância desse construto nos estudos educacionais, tornou-se necessário criar uma escala de medida segundo a qual esta variável assume determinados valores. Para tanto, neste trabalho, optou-se por empregar a Teoria de Resposta ao Item (TRI), por ser a técnica comumente utilizada nas avaliações educacionais desenvolvidas com os microdados do Saeb. Como referencial teórico para a construção desta medida latente, utilizamos o estudo de Alves e Soares (2008).

Para a elaboração do índice, além da escolaridade do pai e da mãe, foram consideradas as seguintes variáveis (ou itens): televisão em cores; vídeo cassete; rádio; geladeira; freezer; máquina de lavar; aspirador de pó; telefone fixo; celular; computador; carro; banheiro; e empregada doméstica. A inclusão desses itens foi baseada na análise de suas curvas características. O índice varia de 0 a 10. Neste caso, quanto mais próximo de 0 (zero), menor é o nível socioeconômico dos alunos e, quanto mais próximo de 10, melhor é sua condição socioeconômica. A utilização da Teoria de Resposta ao Item e a manutenção de itens âncoras ao longo da série histórica garantem a comparabilidade do índice entre os vários anos incluídos no estudo.

A TRI estabelece uma relação probabilística entre a resposta ao item do questionário e o atributo latente para cuja medida contribui.⁸ No caso deste estudo, o atributo latente é a condição socioeconômica familiar e o item refere-se ao conjunto de perguntas sobre a existência ou não de determinados bens duráveis ou de infraestrutura domiciliar. Por exemplo, no caso de um item com respostas graduadas (0 = não possui carro; 1 = possui um carro; 2 = possui dois ou mais carros), a probabilidade de o aluno responder que possui dois ou mais carros em casa será pequena se o nível socioeconômico do aluno for pequeno, ou será grande se o mesmo também o for.

Uma vantagem da TRI em relação aos métodos convencionais é a possibilidade de estimar a medida do NSE mesmo com dados incompletos para algum dos indicadores, pois esse processo é viável mesmo com apenas alguns dos itens respondidos. A ausência de indicadores é muito comum

⁸ Existe uma forma funcional que busca representar a relação entre o atributo latente e a probabilidade de escolha de cada resposta de um item. A curva que caracteriza esta relação é chamada de curva característica do item (CCI) e é ajustada com base em um modelo matemático. No caso deste estudo, como os itens possuem categorias de respostas que podem ser ordenadas entre si, foi utilizado o modelo de respostas graduadas de Samejima (1997). Assim, para cada item, é possível organizar as respostas em ordem crescente, de forma que $x = k$ e $k = 1, 2, \dots, mi$, onde $(mi+1)$ é o número de categorias do i -ésimo item. A probabilidade de um aluno j escolher uma categoria particular ou outra mais alta do item i é representada por um modelo logístico de dois parâmetros (SAMEJIMA, 1997):

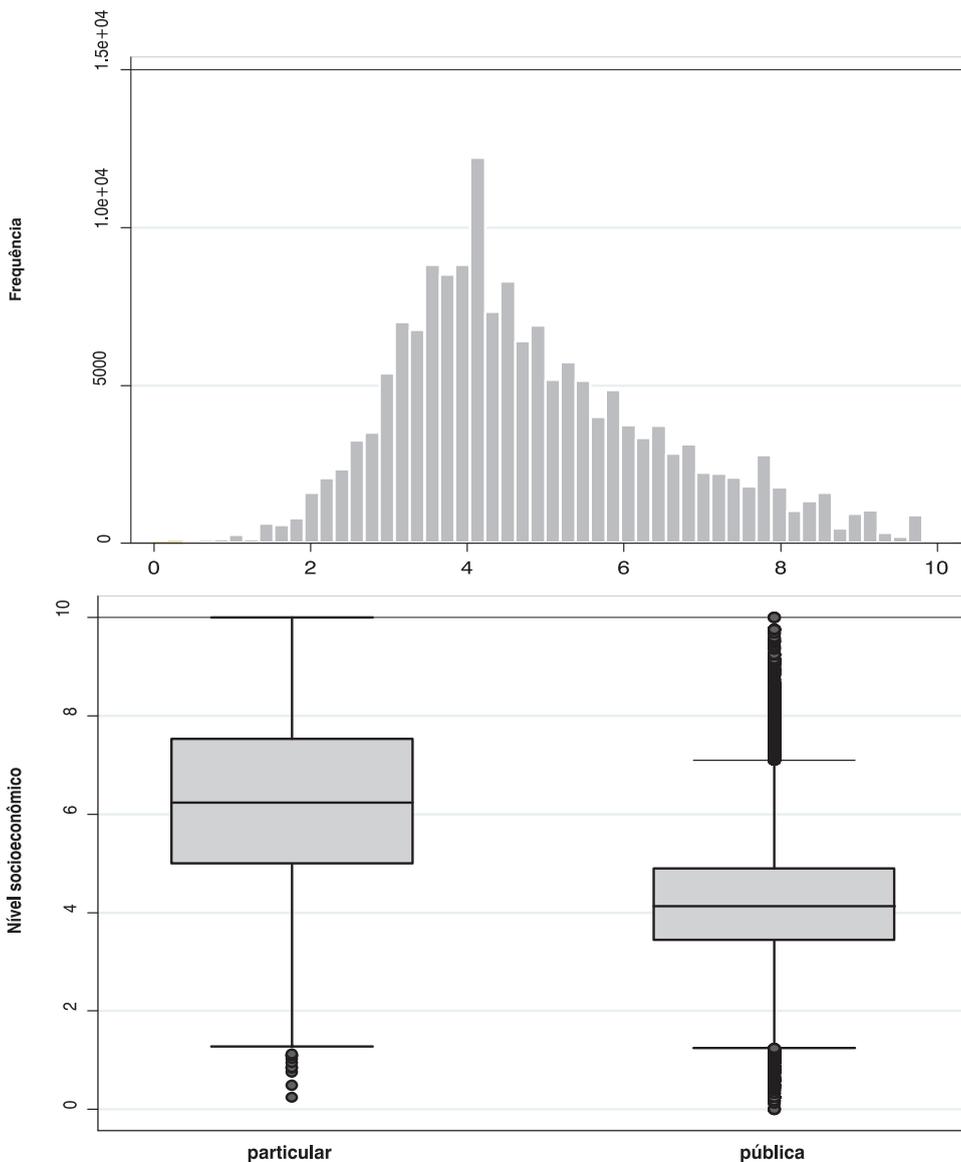
$$P_{i,k}(\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k})}} - \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k+1})}}$$

O atributo que está sendo medido pelo teste é usualmente representado por θ e é normalizado na escala z , onde a média é zero, o desvio-padrão é um, e o conjunto de valores de θ se enquadra numa escala de -3 a 3. Detalhes sobre os resultados deste modelo podem ser encontrados em Rodrigues (2009).

quando são utilizadas as bases de dados do Saeb, uma vez que seus questionários passam por constantes mudanças para aprimoramento das questões a serem investigadas.

O Gráfico 3 apresenta o histograma desta medida para os 164.496 alunos incluídos neste estudo e o nível socioeconômico discriminado por rede de ensino em que esses alunos estão matriculados. Nota-se que há

GRÁFICO 3
Histograma e variação do nível socioeconômico dos alunos da 4ª série do ensino fundamental avaliados em Matemática, segundo redes de ensino Brasil – 1997-2005



Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

maior concentração de alunos com baixo nível socioeconômico (próximo ao valor 4). Isto é esperado, pois aproximadamente 68% dos alunos estão matriculados nas escolas públicas, onde a maior parte provém de famílias com baixo nível socioeconômico. Verifica-se, pelo gráfico de caixas, que os alunos da rede pública de ensino têm, como esperado, condições socioeconômicas mais baixas e mais parecidas entre si.

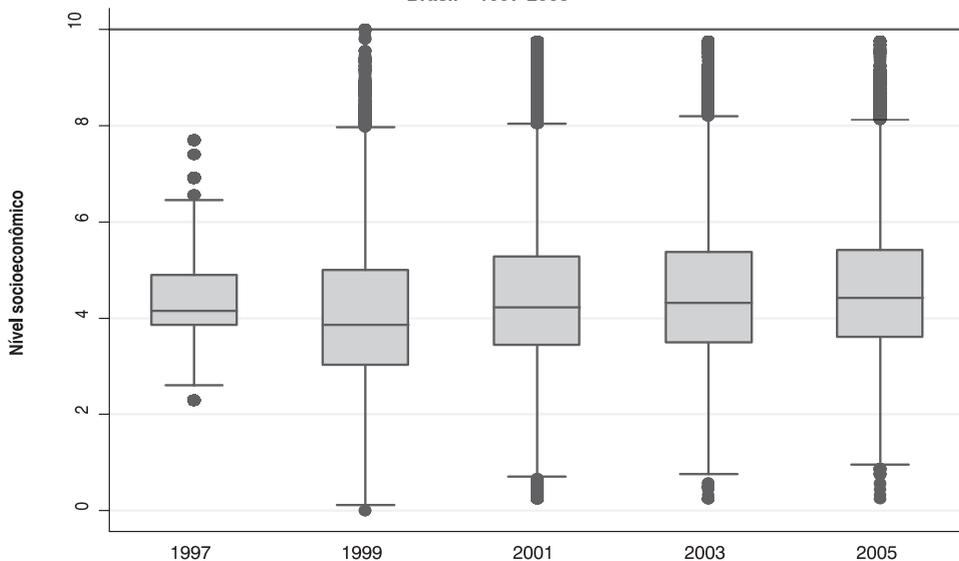
No Gráfico 4, que apresenta a distribuição do nível socioeconômico em cada ciclo do Saeb, observa-se uma redução no nível socioeconômico médio em 1999. A partir de 2001, este índice registra leve aumento e, em 2005, chega a superar o valor atingido em 1997. A melhoria do nível socioeconômico médio dos alunos a partir de 2003 pode ser pensada como uma resposta ao processo de seletividade. Em um primeiro momento, há aumento expressivo nas matrículas de crianças de origem menos favorecida como reflexo das ações políticas de universalização. Uma vez matriculadas no sistema de ensino, estas crianças passarão por maiores dificuldades de aprendizado e, consequentemente, terão maiores chances de serem reprovadas ou abandonarem os estudos.

Tal processo de seleção elevaria a média do nível socioeconômico dos estudantes nos anos posteriores à massificação do ensino.

As taxas de transição (evasão, promoção e repetência) observadas entre 1999 e 2005 colaboram para sustentar esta hipótese. De acordo com o Sistema de Estatísticas Educacionais do Inep (EDUDATABRASIL), a taxa de evasão da 4ª série do ensino fundamental no Brasil elevou-se de 6,0%, em 1999, para 8,1% e 7,4%, em 2003 e 2005, respectivamente. Ao mesmo tempo, a taxa de promoção reduziu-se de 80,9%, em 1999, para 78,3%, em 2003 e 2005, enquanto a taxa de repetência elevou-se de 13,1%, em 1999, para 13,6%, em 2003, e para 14,3%, em 2005, neste mesmo nível de ensino.

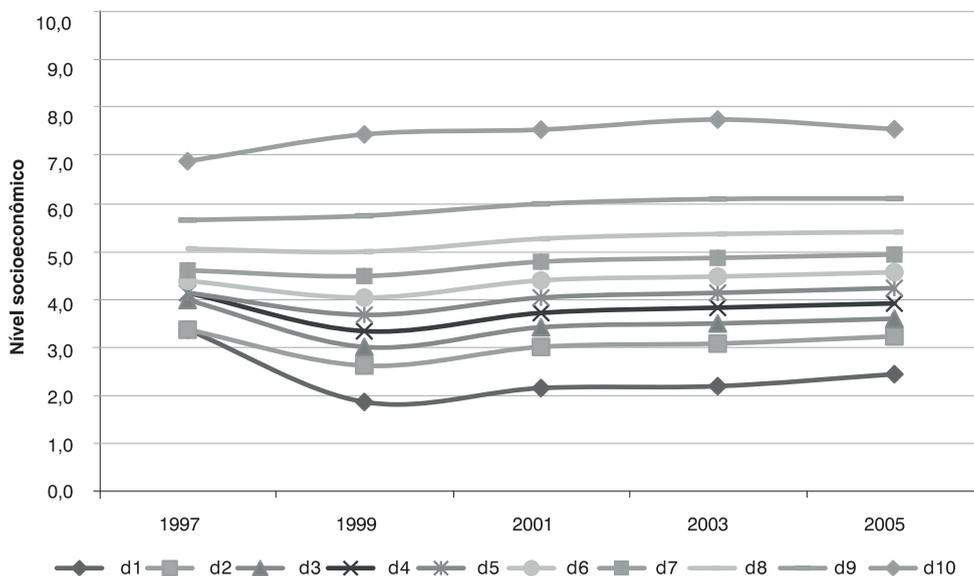
A evolução temporal do nível socioeconômico por decil do desempenho escolar é apresentada no Gráfico 5. Verifica-se que a redução do nível socioeconômico médio em 1999 provém do decréscimo acentuado desta medida entre os alunos com menor desempenho escolar, situados nos primeiros décimos da distribuição. Entre aqueles situados no último décimo, é possível notar uma melhoria em suas condições socioeconômicas.

GRÁFICO 4
Variação do nível socioeconômico dos alunos da 4ª série do ensino fundamental avaliados em Matemática Brasil – 1997-2005



Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

GRÁFICO 5
Evolução do nível socioeconômico dos alunos da 4ª série do ensino fundamental avaliados em Matemática,
segundo decil de desempenho escolar
Brasil – 1997-2005



Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

Variáveis do aluno

A variável-resposta dos modelos de regressão corresponde ao resultado do aluno nos testes padronizados de conhecimento em Matemática. Como covariáveis de caracterização do perfil do aluno, selecionamos as variáveis sociodemográficas (sexo, cor e estrutura familiar), a defasagem idade-série, o nível socioeconômico e a motivação do estudante. A descrição e as formas de medida destas variáveis estão presentes no Quadro 1.

Para cada covariável, incluímos também uma indicadora de dados ausentes, com o intuito de evitar a perda de informações.⁹

Variáveis da escola

Embora seja reconhecida a importância da escola no aprendizado do aluno, ainda há grandes dificuldades em mensurar e identificar os atributos escolares que se associam a este aprendizado. Em geral, atributos importantes, como a didática do professor, não são possíveis de serem coletados pelos questionários contextuais. Por esse motivo, a utilização de dados experimentais (DUFLO, 2001; CHIN, 2002 *apud* GLEWWE; KREMMER, 2006) ou de natureza qualitativa (BANCO MUNDIAL, 2008) aparece como uma opção investigativa em estudos que avaliam a qualidade da educação.

⁹ Nesta análise, o percentual de dados ausentes é elevado e, em vez de ignorá-los, preferimos reparametrizar o modelo ao incluí-los como uma nova categoria. Esta reparametrização apenas cria um novo intercepto para as categorias dos dados ausentes, não gerando nenhum viés sob a hipótese de *missing* completamente aleatório. Sabemos que esta estratégia não resolve o problema de dados faltantes, pois, para tanto, seria necessária uma metodologia que permitisse a correção para não resposta ano a ano, sob a hipótese de *missing* aleatório condicional no vetor de variáveis observáveis, sem afetar a comparatividade intertemporal das decomposições. Estas correções estão fora do escopo deste trabalho.

QUADRO 1
Descrição das variáveis incluídas no modelo do aluno

Variável	Descrição	Forma de medida
Resposta		
PROFIC	Proficiência do aluno	Escala contínua
Covariáveis		
SEXO	Sexo do aluno	Variável indicadora: 0 = Masculino 1 = Feminino
SEXO_A	Indicadora de dado ausente na variável sexo	
MOTIVACAO	Faz lição de casa?	Variável indicadora: 0 = Faz 1 = Não faz
MOTIVACAO_A	Indicadora de dado ausente na variável motivação	
DEFASAGEM	Defasagem idade-série	Variável indicadora: 0 = Não defasado 1 = Defasado
DEFASAGEM_A	Indicadora de dado ausente na variável defasagem	
PARDO	Autodeclaração da cor "parda"	Variável indicadora: 0 = Não pardo 1 = Pardo
PRETO	Autodeclaração da cor "preta"	Variável indicadora: 0 = Não preto 1 = Preto
COR_A	Indicadora de dado ausente na variável cor	
FAMILIA	Reside com ambos os pais?	Variável indicadora: 0 = Reside 1 = Não reside
FAMILIA_A	Indicadora de dado ausente na variável família	
NSE	Nível socioeconômico do aluno	Escala contínua

Tendo em vista que estamos lidando com dados de natureza quantitativa e não-experimentais, como é o caso das informações coletadas pelo Saeb, as variáveis escolares selecionadas neste trabalho são limitadas e podem não refletir adequadamente seu efeito sobre o aprendizado dos alunos. Por exemplo, a qualidade de um professor, medida pela sua escolaridade, pode ser imprecisa, na medida em que não capta o empenho ou a dedicação do docente em sala de aula. Além disso, especificamente neste trabalho, outra limitação surge pelo fato de trabalharmos com vários ciclos do Saeb e, conseqüentemente, da necessidade de encontramos variáveis compatíveis entre eles.

De acordo com as possibilidades presentes nos questionários contextuais do Saeb, optamos por caracterizar a escola por meio da inclusão de variáveis que representam o perfil dos professores e diretores e as condições de infraestrutura escolar. A rede de ensino é usada para controlar as diferenças de qualidade existentes entre os sistemas público e particular de ensino. As diferenças contextuais decorrentes da localização da escola em áreas com maior ou menor desenvolvimento são controladas pelas indicadoras de regiões. A composição do público escolar é controlada pelas características médias dos estudantes, tais como nível socioeconômico, motivação, cor, defasagem idade-série, sexo e estrutura

familiar. O Quadro 2 sintetiza as variáveis selecionadas na esfera escolar.

Estatísticas descritivas

A Tabela 3 apresenta as estatísticas descritivas de todas as variáveis selecionadas para a análise.

Para o cálculo das médias e proporções, foi utilizado o peso amostral referente a cada uma das unidades de análise: aluno e escola. Para os alunos, empregou-se o peso amostral adequado para as análises feitas separadamente para cada uma das disciplinas.

QUADRO 2
Variáveis da escola selecionadas para a análise

Variável	Descrição	Forma de medida
Resposta		
PROFICM	Proficiência média da escola	Escala contínua
Covariáveis		
SEXOM	Proporção de alunos do sexo feminino	Escala contínua
PARDOM	Proporção de alunos pardos	Escala contínua
PRETOM	Proporção de alunos pretos	Escala contínua
FAMILIAM	Proporção de alunos que não residem com ambos os pais	Escala contínua
MOTIVACAOM	Proporção de alunos que fazem lição de casa	Escala contínua
DEFM	Proporção de alunos defasados	Escala contínua
NSEM	Nível socioeconômico médio	Escala contínua
ESTAD	Escola estadual	Variável indicadora: 0 = Não 1 = Sim
MUNIC	Escola municipal	Variável indicadora: 0 = Não 1 = Sim
INFRAEST	Infraestrutura escolar	Escala contínua
EDUCPROF	Escola com mais de 50% dos professores com ensino superior	Variável indicadora: 0 = Não 1 = Sim
EDUCPROF_A	Indicadora de dado ausente na escolaridade do professor	
EXPERPROF	Escola com mais de 60% dos professores em atividades de docência há mais de dez anos	Variável indicadora: 0 = Não 1 = Sim
EXPERPROF_A	Indicadora de dado ausente na experiência do professor	
EDUCDIR	Diretor com ensino superior ou mais	Variável indicadora: 0 = Não 1 = Sim
EDUCDIR_A	Indicadora de dado ausente na escolaridade do diretor	
EXPERDIR	Diretor que trabalha há mais de dez anos em atividades de direção	Variável indicadora: 0 = Não 1 = Sim
EXPERDIR_A	Indicadora de dado ausente na experiência do diretor	
NE	Escola da região Nordeste	Variável indicadora: 0 = Não 1 = Sim
N	Escola da região Norte	Variável indicadora: 0 = Não 1 = Sim
CO	Escola da região Centro-Oeste	Variável indicadora: 0 = Não 1 = Sim
S	Escola da região Sul	Variável indicadora: 0 = Não 1 = Sim
ALUNOS	Número médio de alunos por escola	Escala contínua

TABELA 3
Estatísticas descritivas das escolas e dos alunos da 4ª série do ensino fundamental avaliados em Matemática Brasil – 1997-2005

Variáveis	1997				1999				2001			
	Média/ propor- ção	DP	Mínimo	Máximo	Média/ propor- ção	DP	Mínimo	Máximo	Média/ propor- ção	DP	Mínimo	Máximo
Alunos												
PROFIC: Proficiência do aluno	192,29	43,86	80,37	377,95	182,17	41,00	82,72	355,93	179,17	46,01	59,84	367,25
SEXO: Sexo feminino	0,50	0,50	0	1	0,49	0,50	0	1	0,51	0,50	0	1
SEXO_A: Sem resposta no sexo	0,00	0,04	0	1	0,01	0,10	0	1	0,04	0,19	0	1
PARDO: Aluno que se autodeclara como pardo	0,39	0,49	0	1	0,38	0,48	0	1	0,35	0,48	0	1
PRETO: Aluno que se autodeclara como preto	0,08	0,27	0	1	0,11	0,32	0	1	0,11	0,32	0	1
COR_A: Sem resposta na cor/raça	0,01	0,10	0	1	0,01	0,12	0	1	0,05	0,22	0	1
MOTIVACAO: Não faz dever de casa	0,26	0,44	0	1	0,30	0,46	0	1	0,38	0,49	0	1
MOTIVACAO_A: Sem resposta no dever de casa	0,00	0,06	0	1	0,01	0,10	0	1	0,07	0,25	0	1
FAMILIA: Não reside com ambos os pais	0,14	0,35	0	1	0,15	0,36	0	1	0,20	0,40	0	1
FAMILIA_A: Sem resposta na co-residência	0,02	0,13	0	1	0,03	0,18	0	1	0,09	0,29	0	1
DEFASAGEM: Defasado em relação à idade e série frequentada	0,34	0,47	0	1	0,34	0,47	0	1	0,23	0,42	0	1
DEFASAGEM_A: Sem resposta na defasagem idade-série	0,03	0,17	0	1	0,01	0,12	0	1	0,07	0,25	0	1
NSE: nível socioeconômico do aluno	4,55	1,06	2,30	7,70	4,14	1,59	0,00	10,00	4,45	1,52	0,24	9,76
N (amostra de alunos)	18.588				16.811				50.782			

(continua)

(continuação)

Variáveis	2003				2005			
	Média/ proporção	DP	Mínimo	Máximo	Média/ proporção	DP	Mínimo	Máximo
Alunos								
PROFIC: Proficiência do aluno	180,68	45,01	66,42	369,98	186,09	47,18	65,43	373,44
SEXO: Sexo feminino	0,50	0,50	0	1	0,50	0,50	0	1
SEXO_A: Sem resposta no sexo	0,01	0,07	0	1	0,02	0,13	0	1
PARDO: Aluno que se autodeclara como pardo	0,45	0,50	0	1	0,42	0,49	0	1
PRETO: Aluno que se autodeclara como preto	0,11	0,32	0	1	0,12	0,33	0	1
COR_A: Sem resposta na cor/raça	0,01	0,09	0	1	0,04	0,20	0	1
MOTIVACAO: Não faz dever de casa	0,35	0,48	0	1	0,36	0,48	0	1
MOTIVACAO_A: Sem resposta no dever de casa	0,01	0,09	0	1	0,05	0,23	0	1
FAMILIA: Não reside com ambos os pais	0,06	0,23	0	1	0,09	0,29	0	1
FAMILIA_A: Sem resposta na co-residência	0,01	0,10	0	1	0,04	0,21	0	1
DEFASAGEM: Defasado em relação à idade e série frequentada	0,27	0,44	0	1	0,23	0,42	0	1
DEFASAGEM_A: Sem resposta na defasagem idade-série	0,01	0,07	0	1	0,01	0,11	0	1
NSE: nível socioeconômico do aluno	4,54	1,55	0,24	9,76	4,62	1,45	0,24	9,76
N (amostra de alunos)		40.596				37.719		

(continua)

(continuação)

Variáveis	1997				1999				2001			
	Média/ propor- ção	DP	Mínimo	Máximo	Média/ propor- ção	DP	Mínimo	Máximo	Média/ propor- ção	DP	Mínimo	Máximo
Escolas												
PROFICM: Proficiência média da escola	192,32	25,12	113,36	304,64	184,76	30,02	99,72	307,18	182,60	31,70	99,71	324,15
SEXOM: Proporção de alunos do sexo feminino	0,50	0,09	0,00	1,00	0,49	0,25	0,00	1,00	0,51	0,17	0,00	1,00
PARDOM: Proporção de alunos pardos	0,39	0,15	0,00	1,00	0,38	0,27	0,00	1,00	0,34	0,22	0,00	1,00
PRETOM: Proporção de alunos pretos	0,08	0,08	0,00	0,57	0,11	0,17	0,00	1,00	0,11	0,13	0,00	1,00
MOTIVACAOM: Proporção de alunos que não fazem o dever de casa	0,26	0,10	0,00	0,84	0,29	0,24	0,00	1,00	0,37	0,20	0,00	1,00
FAMILIAM: Proporção de alunos que não residem com ambos os pais	0,14	0,11	0,00	1,00	0,13	0,19	0,00	1,00	0,18	0,18	0,00	1,00
DEFM: Proporção de alunos defasados em relação à idade e série frequentada	0,34	0,22	0,00	1,00	0,34	0,31	0,00	1,00	0,22	0,23	0,00	1,00
NSEM: Nível socioeconômico médio da escola	4,55	0,67	3,30	7,33	4,18	1,39	0,44	9,40	4,54	1,25	1,70	9,11
ESTAD: Escola estadual	0,56	0,50	0,00	1,00	0,35	0,48	0,00	1,00	0,31	0,46	0,00	1,00
MUNIC: Escola municipal	0,31	0,46	0,00	1,00	0,39	0,49	0,00	1,00	0,42	0,49	0,00	1,00
INFRAEST: Infraestrutura escolar	5,77	2,40	0,12	9,88	6,00	0,75	2,85	7,15	6,79	2,12	0,08	9,92
EDUCPROF: Escola onde 50% ou mais dos professores têm curso superior	0,34	0,47	0	1	0,29	0,45	0	1	0,34	0,47	0	1
EDUCPROF_A: Escola com dado ausente na escolaridade do professor	0,01	0,07	0	1	0,13	0,33	0	1	0,08	0,28	0	1

(continua)

(continuação)

Variáveis	2003				2005			
	Média/ proporção	DP	Mínimo	Máximo	Média/ proporção	DP	Mínimo	Máximo
Escolas								
PROFICM: Proficiência média da escola	184,95	31,98	99,21	342,21	188,09	30,94	108,97	296,78
SEXOM: Proporção de alunos do sexo feminino	0,50	0,17	0,00	1,00	0,50	0,15	0,00	1,00
PARDOM: Proporção de alunos pardos	0,44	0,25	0,00	1,00	0,41	0,22	0,00	1,00
PRETOM: Proporção de alunos pretos	0,10	0,13	0,00	1,00	0,12	0,12	0,00	1,00
MOTIVACAOM: Proporção de alunos que não fazem o dever de casa	0,34	0,19	0,00	1,00	0,36	0,17	0,00	1,00
FAMILIAM: Proporção de alunos que não residem com ambos os pais	0,05	0,09	0,00	1,00	0,09	0,10	0,00	0,71
DEFM: Proporção de alunos defasados em relação à idade e série frequentada	0,26	0,25	0,00	1,00	0,22	0,21	0,00	1,00
NSEM: Nível socioeconômico médio da escola	4,66	1,37	1,13	9,24	4,67	1,15	1,95	9,20
ESTAD: Escola estadual	0,28	0,45	0,00	1,00	0,33	0,47	0,00	1,00
MUNIC: Escola municipal	0,45	0,50	0,00	1,00	0,49	0,50	0,00	1,00
INFRAEST: Infraestrutura escolar	6,67	2,12	0,08	9,92	6,66	2,03	0,08	9,92
EDUCPROF: Escola onde 50% ou mais dos professores têm curso superior	0,60	0,49	0	1	0,66	0,47	0	1
EDUCPROF_A: Escola com dado ausente na escolaridade do professor	0,01	0,10	0	1	0,09	0,29	0	1

Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

Método de decomposição baseado em simulações contrafactuais

Para cumprir o objetivo deste estudo, que é identificar e quantificar o efeito de mudanças ocorridas na “quantidade” e “retorno (coeficientes)” dos fatores associados ao desempenho escolar (em especial, o nível socioeconômico) sobre as mudanças no desempenho no período de 1997 a 2005, foram utilizadas técnicas empíricas de decomposição baseadas em simulações contrafactuais.

Na literatura educacional, a relação entre o desempenho escolar e os fatores associados é comumente modelada por meio de uma função de produção educacional. Como resultados desta função, temos o grau de sensibilidade do desempenho escolar aos insumos individuais e escolares inseridos neste processo (representado pela estrutura de coeficientes da regressão) e o componente não explicado (o resíduo), que, em geral, é atribuído às características não observáveis pelo pesquisador, mas que exercem influência sobre o aprendizado do aluno.

O uso das técnicas de decomposição baseadas em simulações contrafactuais permite medir o impacto da variação em cada um dos componentes da função de produção educacional (insumos, coeficientes e resíduo) sobre as variações intertemporais ocorridas no desempenho escolar (entendido como produto desta função). Neste estudo, utilizamos o método de decomposição elaborado por Juhn, Murphy e Pierce (1993), que permite, além da decomposição do resultado médio de uma regressão, decompor o resultado nos quantis e em medidas de desigualdade. Tal metodologia ainda incorpora à decomposição o termo de erro, que supostamente é nulo na média, mas diferente de zero nos quantis. Como em uma função de produção educacional o ajuste é geralmente baixo, o erro tem papel importante, pois responde pelas características não

observáveis. Assim, tornou-se interessante utilizar uma abordagem econométrica que fosse capaz de levar em conta as diferenças em suas distribuições.

Para lidar com a hierarquia presente nos dados educacionais e, ao mesmo tempo, decompor a variação do desempenho nas unidades micro (aluno) e macro (escola), adotamos a estratégia de implementar as regressões separadamente em cada uma dessas unidades de análise.¹⁰

Alunos como unidade de análise

Para formalizarmos o método, supomos que a proficiência escolar do indivíduo i e j , em dois períodos, t_0 e t_1 , respectivamente, possa ser modelada pelas seguintes equações:

$$Y_{it_0} = X_{it_0} \beta_{it_0} + u_{it_0} \quad (1)$$

$$Y_{jt_1} = X_{jt_1} b_{jt_1} + u_{jt_1} \quad (2)$$

A equação (1) indica que o desempenho escolar do aluno i no ano t_0 está associado a um vetor de características individuais observáveis, X_{it_0} , a uma constante β_{it_0} e à heterogeneidade não observada sintetizada no termo de erro, u_{it_0} , para o qual assume-se média condicional nula ($E[u_{it_0} / X_{it_0}] = 0$). Esta mesma formalização pode ser mantida para um aluno j avaliado em um período posterior, t_1 , como mostra a equação (2).

Tomando como base estas duas equações e seguindo a especificação de Juhn et al. (1993), pode-se dividir o termo de erro em dois componentes: o percentil do indivíduo i , θ_{it_0} , na distribuição do erro; e a função de distribuição acumulada do erro, F_{it_0} . Por definição, tem-se a função de distribuição acumulada condicional do erro definida como:

$$u_{it_0} = F_{it_0}^{-1}(\theta_{it_0} / X_{it_0}) \quad (3),$$

onde $F_{it_0}^{-1}(\cdot / X)$ é a função inversa da distribuição do erro acumulada condicional

¹⁰ É importante esclarecer que o objetivo inicial era fazer a decomposição utilizando o modelo hierárquico linear. Porém, chegamos à conclusão de que a estratégia adotada aqui (modelo de regressão linear clássico agregado para a escola) tornou-se mais prática devido à inexistência de um método, bem como de rotinas computacionais, que decomponha, simultaneamente, ambos os níveis do modelo hierárquico. Ao comparar os coeficientes em ambos os métodos (ver tabela do Anexo), percebemos que eles são muito similares e isto valida os resultados da decomposição baseada na regressão linear clássica. Ademais, a opção pela utilização da escola como unidade de análise, em vez do aluno, se justifica pela possibilidade de identificarmos o resíduo pertencente à escola.

em X . Assim, as equações (1) e (2) podem ser reescritas como se segue:

$$Y_{it_0} = X_{it_0} \beta_{t_0} + F_{t_0}^{-1} (\theta_{it_0} / X_{it_0}) \quad (4)$$

$$Y_{jt_1} = X_{jt_1} \beta_{t_1} + F_{t_1}^{-1} (\theta_{jt_1} / X_{jt_1}) \quad (5)$$

Para realizar a análise contrafactual, utilizamos a equação do período t_0 (eq. 4) como referência. Em seguida, empregamos os coeficientes e a distribuição dos erros desta equação para criar as distribuições contrafactualis da proficiência escolar em t_1 e, assim, isolar o efeito de mudanças nas características, nos coeficientes e no resíduo entre dois pontos no tempo.¹¹

A primeira distribuição contrafactual, $Y_{jt_0}^1$, é criada quando se utilizam os coeficientes e a distribuição dos erros em t_0 , mantendo fixa apenas a distribuição dos atributos individuais em t_1 :

$$Y_{jt_0}^1 = X_{jt_1} \beta_{t_0} + F_{t_0}^{-1} (\theta_{jt_1} / X_{jt_1}) \quad (6)$$

Pela equação (6), percebe-se que a variação da proficiência entre t_0 e t_1 só ocorrerá se a distribuição das características individuais variarem temporalmente.

Da mesma forma, é possível calcular a segunda distribuição contrafactual do desempenho escolar no ano t_1 variando as características individuais e os retornos a estas características, mantendo constante apenas a distribuição dos erros:

$$Y_{jt_1}^2 = X_{jt_1} \beta_{t_1} + F_{t_1}^{-1} (\theta_{jt_1} / X_{jt_1}) \quad (7)$$

Com base nessas duas distribuições fictícias, o cálculo dos efeitos composição (C), retorno (R) e resíduo (Rs) entre t_0 e t_1 pode ser facilmente implementado da seguinte forma:

$$C = Y_{jt_1}^1 - Y_{it_0} = [X_{jt_1} \beta_{t_0} + F_{t_0}^{-1} (\theta_{jt_1} / X_{jt_1})] - [X_{it_0} \beta_{t_0} + F_{t_0}^{-1} (\theta_{it_0} / X_{it_0})] \quad (8)$$

$$R = Y_{jt_1}^2 - Y_{jt_1}^1 = [X_{jt_1} \beta_{t_1} + F_{t_1}^{-1} (\theta_{jt_1} / X_{jt_1})] - [X_{jt_1} \beta_{t_0} + F_{t_0}^{-1} (\theta_{jt_1} / X_{jt_1})] \quad (9)$$

$$Rs = Y_{jt_1}^2 - Y_{jt_0}^1 = [X_{jt_1} \beta_{t_1} + F_{t_1}^{-1} (\theta_{jt_1} / X_{jt_1})] - [X_{jt_1} \beta_{t_0} + F_{t_0}^{-1} (\theta_{jt_1} / X_{jt_1})] \quad (10)$$

Observa-se que na equação (8) apenas as características (Xs) variam entre t_0 e t_1 , o

que permite interpretá-las como o efeito de mudanças na composição dos atributos dos alunos entre esses dois pontos no tempo. De forma similar, a equação (9) pode ser atribuída ao efeito retorno, pois apenas mudanças nos coeficientes são consideradas entre t_0 e t_1 . Por último, a equação (10) capta apenas mudanças nos componentes não observáveis, que, por sua vez, podem explicar as diferenças na distribuição do desempenho escolar entre o período de referência, t_0 , e de comparação, t_1 .

Escolas como unidade de análise

Para identificar os efeitos composição, retorno e resíduo no nível macro, definido pelas unidades escolares, utilizamos apenas a variação entre as escolas. Neste caso, adotou-se a estratégia convencional da regressão na média como variável-resposta, conforme o seguinte modelo:¹²

$$\bar{Y}_{jt_0} = \gamma_{00t_0} + \gamma_{01t_0} Z_{jt_0} + \gamma_{10t_0} \bar{X}_{jt_0} + \bar{\eta}_{jt_0} \quad (11)$$

Neste modelo, \bar{Y}_{jt_0} representa a proficiência média dos alunos da escola j no ano t_0 ; Z_{jt_0} corresponde às variáveis da escola j no ano t_0 (incluindo as variáveis dos professores, diretores, infraestrutura escolar, rede de ensino e localização geográfica); \bar{X}_{jt_0} refere-se à média das características individuais dos alunos da escola j no ano t_0 ; e $\bar{\eta}_{jt_0}$ representa o termo de erro da escola j no ano t_0 . Formalmente, estas variáveis são definidas como:

$$\bar{Y}_{jt_0} = \frac{1}{M_j} \sum_{i=1}^{M_j} Y_{ijt_0} \quad (12)$$

$$\bar{X}_{jt_0} = \frac{1}{M_j} \sum_{j=1}^M X_{ijt_0} \quad (13)$$

$$\bar{\eta}_{jt_0} = \frac{1}{M_j} \sum_{i=1}^{M_j} r_{ijt_0} + u_{jt_0} \quad (14)$$

Sendo M_j o número de alunos na escola j .

Partindo da equação (11), assumimos que $\bar{\eta}_{jt_0}$ tem uma função de distribuição G_j

¹¹ Para efetuar as decomposições nos percentis, são utilizados os coeficientes da regressão na média. Neste caso, estamos assumindo que todos os alunos têm o mesmo retorno em suas características. No que se refere ao efeito de mudanças nas características, são considerados os valores médios ou as proporções das características (Xs) nos percentis.

¹² Utilizamos o método dos Mínimos Quadrados Ordinários para estimar os parâmetros. Como estamos particularmente interessados em coeficientes não tendenciosos e consistentes, mas não necessariamente eficientes, consideramos que o estimador de mínimos quadrados ordinários garante estas propriedades.

e que q_{jt} representa o percentil da escola j nesta distribuição:

$$\bar{\eta}_{jt} = G_t^{-1}(q_{jt} | Z_{jt}, \bar{X}_{jt}) \tag{15}$$

Assim, é possível decompor a variação na proficiência média da escola nos três componentes propostos por Juhn et al. (1993), entre dois períodos: t_0 (referência) e t_1 (comparação). Supomos que a distribuição da proficiência média da escola j no ano t_1 possa ser reconstruída como:

$$\bar{Y}_{jt_1}^1 = \gamma_{00t_0} + \gamma_{01t_0} Z_{jt_1} + \gamma_{10t_0} \bar{X}_{jt_1} + G_{t_0}^{-1}(q_{jt_1} | Z_{jt_1}, \bar{X}_{jt_1}) \tag{16}$$

$$\bar{Y}_{jt_1}^2 = \gamma_{00t_1} + \gamma_{01t_1} Z_{jt_1} + \gamma_{10t_1} \bar{X}_{jt_1} + G_{t_0}^{-1}(q_{jt_1} | Z_{jt_1}, \bar{X}_{jt_1}) \tag{17}$$

$$\bar{Y}_{jt_1}^3 = \gamma_{00t_1} + \gamma_{01t_1} Z_{jt_1} + \gamma_{10t_1} \bar{X}_{jt_1} + G_{t_1}^{-1}(q_{jt_1} | Z_{jt_1}, \bar{X}_{jt_1}) \tag{18}$$

Com base nestas três distribuições, o cálculo dos efeitos composição (C), retorno (R) e resíduo (Rs) entre os períodos t_0 e t_1 pode ser feito da seguinte forma:

$$C = \bar{Y}_{jt_1}^1 - \bar{Y}_{jt_0}^1 = [\gamma_{00t_0} + \gamma_{01t_0} Z_{jt_1} + \gamma_{10t_0} \bar{X}_{jt_1} + G_{t_0}^{-1}(q_{jt_1} | Z_{jt_1}, \bar{X}_{jt_1})] - [\gamma_{00t_0} + \gamma_{01t_0} Z_{jt_0} + \gamma_{10t_0} \bar{X}_{jt_0} + G_{t_0}^{-1}(q_{jt_0} | Z_{jt_0}, \bar{X}_{jt_0})] \tag{19}$$

$$R = \bar{Y}_{jt_1}^2 - \bar{Y}_{jt_1}^1 = [\gamma_{00t_1} + \gamma_{01t_1} Z_{jt_1} + \gamma_{10t_1} \bar{X}_{jt_1} + G_{t_0}^{-1}(q_{jt_1} | Z_{jt_1}, \bar{X}_{jt_1})] - [\gamma_{00t_0} + \gamma_{01t_0} Z_{jt_1} + \gamma_{10t_0} \bar{X}_{jt_1} + G_{t_0}^{-1}(q_{jt_1} | Z_{jt_1}, \bar{X}_{jt_1})] \tag{20}$$

$$Rs = \bar{Y}_{jt_1}^3 - \bar{Y}_{jt_1}^2 = [\gamma_{00t_1} + \gamma_{01t_1} Z_{jt_1} + \gamma_{10t_1} \bar{X}_{jt_1} + G_{t_1}^{-1}(q_{jt_1} | Z_{jt_1}, \bar{X}_{jt_1})] - [\gamma_{00t_1} + \gamma_{01t_1} Z_{jt_1} + \gamma_{10t_1} \bar{X}_{jt_1} + G_{t_0}^{-1}(q_{jt_1} | Z_{jt_1}, \bar{X}_{jt_1})] \tag{21}$$

Com as equações (19), (20) e (21), torna-se possível estimar o efeito de variações nas características das escolas (efeito composição escolar), na taxa de retorno destas características (efeito retorno escolar) e na heterogeneidade não observada da escola (efeito resíduo escolar) sobre as variações na proficiência média da escola ao longo dos anos.

A importância do nível socioeconômico para as variações no desempenho escolar

Resultados da decomposição na média

A Tabela 4 apresenta os resultados da decomposição da diferença de médias do desempenho escolar entre o ano de referência

(1997) e os anos de comparação (1999, 2001, 2003 e 2005), para os exercícios em que o aluno foi utilizado como unidade de análise.

Observa-se que, entre 1997 e 1999, período em que houve acentuado declínio no desempenho escolar médio dos alunos (queda de 10,12 pontos ou -0,23 em múltiplos de desvio-padrão), as mudanças composicionais do alunado (ou o efeito composição) foram responsáveis por 50% da diferença total no desempenho escolar médio. Entre as características dos alunos incluídas na análise, o nível socioeconômico se destaca como aquela que mais contribuiu para este resultado. A redução do nível socioeconômico médio dos alunos, em 1999, respondeu pelo declínio de 4,5 pontos do desempenho médio, o que representa 89% do efeito composição total e 44% da diferença de médias.

As demais características que compõem o efeito composição possuem pequena contribuição para o diferencial de médias observado entre os períodos.

No que se refere ao efeito retorno, ou seja, no efeito de mudanças temporais na magnitude dos coeficientes das características dos alunos, constatamos, de maneira geral, que o mesmo possui maior relevância para explicar o decréscimo do desempenho escolar médio ao longo dos anos. Em alguns períodos de comparação, como em 1997-2003 e 1997-2005, o efeito retorno total é capaz de explicar a totalidade do hiato entre as médias. Caso não houvesse resultados compensatórios, o efeito retorno total em 2003 e 2005 teria reduzido o desempenho escolar médio em 13,32 pontos (-0,30 em múltiplos de desvio-padrão) e 7,93 pontos (-0,18 em múltiplos de desvio-padrão), respectivamente.

Analisando separadamente o efeito retorno de cada uma das características individuais e familiares, percebe-se a elevada magnitude do efeito retorno do nível socioeconômico e sua contribuição para redução das médias ao longo dos anos. Isoladamente, o efeito retorno associado às mudanças no coeficiente do nível socioeconômico seria capaz de diminuir em 16,50 (-0,37 em múltiplos de desvio-padrão), 11,69 (-0,26 em múltiplos de desvio-padrão), 4,27

(-0,1 em múltiplos de desvio-padrão) e 4,49 pontos (-0,1 em múltiplos de desvio-padrão) o desempenho escolar médio em 1999, 2001, 2003 e 2005, respectivamente.

Além da variação do coeficiente do nível socioeconômico, verifica-se que o efeito da variação do intercepto sobre a variação no desempenho médio também

se destaca entre os demais. O intercepto pode ser interpretado como a qualidade média das escolas. Em um primeiro momento, entre 1999 e 2001, o intercepto atua no sentido de elevar o desempenho médio global. Posteriormente, entre 2003 e 2005, o mesmo age no sentido de reduzir a média global.

TABELA 4
Decomposição da variação do desempenho escolar na média, para alunos da 4ª série do ensino fundamental avaliados em Matemática
Brasil – 1997-2005 (unidade de análise: aluno)

Componentes da decomposição	Ciclos do Saeb							
	1997-1999		1997-2001		1997-2003		1997-2005	
A: Desempenho médio do ano de comparação	182,17		179,17		180,68		186,09	
B: Desempenho médio do ano de referência (1997)	192,29		192,29		192,29		192,29	
	%		%		%		%	
Diferença total (A - B)	-10,12	100,00	-13,12	100,00	-11,61	100,00	-6,2	100,00
Efeito composição: $\sum (\bar{X}_{it} - \bar{X}_{t0}) \beta_{t0}$	-5,08	50,20	-2,58	19,69	1,70	-14,66	1,73	-27,83
Sexo: feminino	0,04	-0,40	-0,04	0,29	0,02	-0,17	0,02	-0,32
Cor: pardo	0,02	-0,23	0,07	-0,53	-0,09	0,81	-0,05	0,81
Cor: preto	-0,43	4,26	-0,44	3,38	-0,42	3,58	-0,54	8,72
Família: não reside com ambos os pais	-0,24	2,41	-0,69	5,22	-0,52	4,51	-0,60	9,67
Motivação: não faz lição de casa	-0,16	1,57	-0,97	7,41	1,31	-11,30	0,81	-13,05
Defasagem escolar	-0,07	0,65	2,14	-16,35	1,27	-10,95	2,10	-33,87
Nível socioeconômico	-4,50	44,44	-1,12	8,56	-0,08	0,71	0,69	-11,18
Dados ausentes	0,25	-2,49	-1,54	11,70	0,21	-1,85	-0,71	11,39
Efeito retorno: $\sum (\beta_{it} - \beta_{t0}) \bar{X}_{it}$	-5,04	49,82	-10,54	80,33	-13,32	114,72	-7,93	127,92
Sexo: feminino	1,86	-18,36	0,20	-1,55	0,50	-4,32	1,29	-20,75
Cor: pardo	0,15	-1,49	-0,41	3,10	0,31	-2,70	-1,33	21,39
Cor: preto	-0,37	3,69	0,02	-0,18	-0,10	0,84	-0,60	9,65
Família: não reside com ambos os pais	1,88	-18,55	0,45	-3,40	1,69	-14,55	1,92	-31,00
Motivação: não faz lição de casa	0,55	-5,40	-1,62	12,37	0,08	-0,68	-0,67	10,73
Defasagem escolar	1,64	-16,20	0,65	-4,93	0,73	-6,31	0,75	-12,07
Nível socioeconômico	-16,50	163,02	-11,69	89,13	-4,27	36,77	-4,49	72,49
Intercepto	6,09	-60,16	3,58	-27,29	-12,06	103,83	-4,91	79,15
Dados ausentes	-0,33	3,28	-1,72	13,10	-0,21	1,83	0,10	-1,67

Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

Nota: Os valores em % indicam a contribuição de cada um dos componentes para a diferença total no desempenho escolar entre os períodos.

A diferença é calculada subtraindo-se do desempenho médio do ano de comparação o desempenho médio do ano de referência.

Na literatura de avaliação educacional, o intercepto (desempenho médio das escolas) é considerado aleatório e, portanto, possui certa variabilidade. Estudos mostram que cerca de 26% a 50% da variação do desempenho dos alunos provém de diferenças no desempenho entre as escolas.¹³ Devido à elevada magnitude da variância explicada pela escola, consideramos importante abrir

a “caixa-preta” do intercepto e realizar as decomposições utilizando a escola como unidade de análise. O desempenho médio da escola é importante porque reflete as políticas educacionais voltadas para aumentar a qualidade do ensino. Os resultados são apresentados na Tabela 5.

É possível notar que os resultados são similares aos que foram encontrados para

TABELA 5
Decomposição da variação do desempenho escolar na média, para alunos da 4ª série do ensino fundamental avaliados em Matemática
Brasil – 1997-2005 (unidade de análise: escola)

Componentes da decomposição	Ciclos do Saeb							
	1997-1999		1997-2001		1997-2003		1997-2005	
A: Desempenho médio do ano de comparação (t1)	184,76		182,60		184,95		188,09	
B: Desempenho médio do ano de referência (t0 = 1997)	192,32		192,32		192,32		192,32	
	%		%		%		%	
Diferença total (A - B)	-7,56	100,00	-9,72	100,00	-7,37	100,00	-4,23	100,00
Efeito composição total	-7,58	100,30	-0,34	3,48	5,97	-81,01	4,43	-104,78
	$\sum (\bar{X}_{it} - \bar{X}_{i0}) \beta_{i0}$							
Proporção de alunos do sexo feminino	0,10	-1,37	-0,02	0,21	0,02	-0,31	0,01	-0,14
Proporção de alunos pardos	0,16	-2,16	0,52	-5,33	-0,43	5,89	-0,18	4,29
Proporção de alunos pretos	-0,66	8,71	-0,54	5,60	-0,48	6,45	-0,72	16,93
Proporção de alunos que não residem com ambos os pais	-0,09	1,13	-0,29	2,99	-0,21	2,90	-0,25	5,88
Proporção de alunos que não fazem lição de casa	0,24	-3,22	-1,37	14,10	3,42	-46,34	1,96	-46,37
Proporção de alunos com defasagem idade-série	-0,06	0,79	1,34	-13,79	0,84	-11,35	1,32	-31,32
Nível socioeconômico médio	-6,61	87,44	-0,21	2,19	1,94	-26,29	2,07	-48,89
Escola estadual	1,05	-13,87	1,25	-12,84	1,41	-19,12	1,15	-27,28
Escola municipal	-0,54	7,19	-0,73	7,52	-0,90	12,24	-1,16	27,39
Infraestrutura escolar	0,14	-1,83	0,61	-6,26	0,53	-7,25	0,53	-12,47
Qualidade do professor	-0,12	1,57	0,31	-3,24	0,86	-11,67	1,04	-24,65
Qualidade do diretor	0,19	-2,50	0,01	-0,09	0,02	-0,30	-0,56	13,28
Escola da Região Nordeste	-0,93	12,33	-0,92	9,43	-0,89	12,00	-0,35	8,17
Escola da Região Norte	0,06	-0,78	0,06	-0,57	0,07	-1,00	0,01	-0,24
Escola da Região Centro-Oeste	-0,02	0,31	-0,02	0,25	-0,02	0,23	-0,01	0,15
Escola da Região Sul	-0,14	1,88	-0,10	1,03	-0,11	1,51	-0,11	2,56
Dados ausentes	-0,35	4,69	-0,22	2,28	-0,10	1,39	-0,34	7,93

(continua)

¹³ O percentual da variância do desempenho escolar atribuído à diferença entre as escolas depende da série, ano e disciplina avaliados pelo Saeb. É importante deixar claro que este elevado percentual provém dos resultados dos modelos nulos ou incondicionais. Não há, portanto, controle pelo nível socioeconômico médio dos alunos da escola. Ver Fletcher (1998), Ferrão e Fernandes (2001), Pichetti et al. (2008).

(continuação)

Componentes da decomposição	Ciclos do Saeb							
	1997-1999		1997-2001		1997-2003		1997-2005	
Efeito retorno: $\sum (\beta_{it} - \beta_{i0}) \bar{X}_{it}$	0,02	-0,30	-9,38	96,53	-13,34	181,01	-8,66	204,78
Proporção de alunos do sexo feminino	1,77	-23,46	1,94	-19,99	1,99	-27,02	4,44	-104,98
Proporção de alunos pardos	2,96	-39,20	3,73	-38,36	1,54	-20,96	3,46	-81,87
Proporção de alunos pretos	-0,71	9,41	0,76	-7,85	0,86	-11,60	0,34	-7,94
Proporção de alunos que não residem com ambos os pais	0,32	-4,30	-3,63	37,32	1,45	-19,65	-1,15	27,21
Proporção de alunos que não fazem lição de casa	2,22	-29,34	-1,63	16,77	-0,41	5,60	-0,36	8,52
Proporção de alunos com defasagem idade-série	2,07	-27,43	1,17	-12,01	1,85	-25,15	-0,08	1,98
Nível socioeconômico médio	-45,96	608,03	-27,93	287,33	-33,76	457,97	-36,31	858,57
Escola estadual	-4,39	58,06	-1,53	15,71	-2,72	36,94	-3,81	90,14
Escola municipal	-3,43	45,34	-2,06	21,19	-4,89	66,32	-5,51	130,31
Infraestrutura escolar	11,06	-146,37	1,65	-16,96	1,20	-16,27	-4,55	107,59
Qualidade do professor	-0,99	13,10	-1,58	16,23	-1,13	15,30	-1,32	31,21
Qualidade do diretor	0,73	-9,67	1,27	-13,03	0,09	-1,15	1,00	-23,67
Escola da Região Nordeste	0,73	-9,64	1,35	-13,89	0,07	-0,89	-1,74	41,16
Escola da Região Norte	0,46	-6,05	0,28	-2,90	0,09	-1,19	0,00	0,05
Escola da Região Centro-Oeste	0,07	-0,96	0,00	-0,03	-0,32	4,31	-0,43	10,14
Escola da Região Sul	0,40	-5,27	0,49	-4,99	0,35	-4,70	0,57	-13,52
Intercepto	32,62	-431,55	16,12	-165,86	20,54	-278,62	37,04	-875,70
Dados ausentes	0,07	-0,98	0,21	-2,17	-0,13	1,77	-0,24	5,60

Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

Nota: Os valores em % indicam a contribuição de cada um dos componentes para a diferença total no desempenho escolar entre os períodos.

A diferença é calculada subtraindo-se do desempenho médio do ano de comparação o desempenho médio do ano de referência.

a unidade micro (alunos), em termos da direção dos sinais dos efeitos composição e retorno. O efeito composição total aparece com sinal negativo nos dois primeiros períodos de comparação (1997-1999 e 1997-2001), embora neste segundo seja praticamente inexistente. Nos dois últimos períodos (1997-2003 e 1997-2005), o sinal do efeito composição total passa a ser positivo, contribuindo para elevar o desempenho médio das escolas.

Observamos que a principal contribuição para o efeito composição total provém da variação na composição do nível socioeconômico médio dos alunos. Entre 1997 e 1999, as mudanças na composição do alunado em termos de suas origens sociais, por si só, foram responsáveis por um declínio de 6,61 pontos (-0,15 em múltiplos

de desvio-padrão) no desempenho médio das escolas, ou seja, 87,44% do declínio total das médias observado neste período. Este é um resultado esperado, pois 1999 foi um ano marcado pela expansão do acesso ao ensino, principalmente pela entrada de crianças e jovens provenientes de famílias mais pobres. Neste caso, o aumento na participação de alunos com maiores dificuldades de aprendizado afetou diretamente o desempenho médio dos sistemas de ensino.

A análise do efeito retorno mostra que o nível socioeconômico médio das escolas destaca-se como o principal componente. Este efeito, por si só, seria capaz de explicar 100% do declínio total do desempenho médio das escolas e, além disso, caso não houvesse efeitos compensatórios de outras fontes, seria capaz de reduzir em

45,96 pontos o desempenho médio das escolas entre 1997 e 1999, o que corresponde a aproximadamente um declínio de 1 desvio-padrão. A elevada magnitude do efeito retorno do nível socioeconômico médio da escola está associada à redução da sensibilidade do desempenho escolar ao nível socioeconômico médio (isto é, ao decréscimo dos coeficientes atrelados ao nível socioeconômico médio).

É importante considerar que os atributos médios dos alunos de determinada escola (a variável agregada) não necessariamente têm o mesmo significado que a variável individual e, portanto, podem medir aspectos distintos. O nível socioeconômico do aluno no estudo micro está relacionado aos recursos econômicos ou ao poder aquisitivo da família do aluno. No estudo macro (escola), o nível socioeconômico está medindo a “bagagem” cultural dos alunos de determinada escola e é chamado pela literatura de efeito dos pares. Neste caso, aventamos a hipótese da existência de um efeito de pares negativo, ou seja, a entrada de alunos com maiores dificuldades de aprendizado pode ter afetado negativamente o aprendizado dos alunos com elevado desempenho potencial e melhores condições socioeconômicas. Este seria um efeito indireto do aumento da cobertura escolar.

Resultados da decomposição nas medidas de diferença entre os percentis: 90-10, 90-50 e 50-10

Na análise do conjunto de alunos em todo o Brasil, é natural haver diferenças em seus resultados escolares, pois muitas são as características que diferenciam um aluno de outro: habilidade inata; escolhas individuais; características adscritas (como raça/cor e sexo); ambiente familiar; etc. O mesmo ocorre em relação ao conjunto de escolas, já que estas também se diferenciam em vários aspectos, como qualificação do diretor e dos docentes, proposta pedagógica, tamanho das turmas, critério de formação das turmas, composição socioeconômica do alunado, entre outros.

Portanto, as diferenças de resultados entre alunos e entre escolas são comuns

e poderiam existir mesmo em contexto de igualdade, ou seja, mesmo se as características que os diferenciam não os colocassem em desvantagem em termos do aprendizado adquirido. Portanto, a intenção em quantificar e decompor as diferenças entre os percentis é mostrar não a existência deste hiato, mas sim se houve aumento ou redução da distância entre os resultados alcançados pelos alunos e pelas escolas em diferentes partes da distribuição ao longo dos anos. Pretendemos mostrar, ainda, em que medida este aumento ou redução no hiato pode ser atribuído às mudanças nas características médias (efeito composição), no retorno a estas características (efeito retorno) e na heterogeneidade não observada (efeito resíduo).

A Tabela 6 apresenta os resultados da decomposição nas medidas de diferença entre os percentis (90-10, 90-50, 50-10), para os alunos e as escolas, nos períodos analisados. A primeira coluna traz a diferença total nas medidas de diferença do desempenho escolar entre os percentis entre dois pontos no tempo. Na segunda, terceira e quarta colunas tem-se a contribuição dos efeitos composição, retorno e resíduo. A soma desses três componentes equivale ao valor da primeira coluna. As duas últimas colunas mostram a contribuição isolada dos efeitos composição e retorno do nível socioeconômico dos alunos e das escolas.

Alunos como unidade de análise

No período 1997-1999, verifica-se redução no hiato entre os alunos com melhor desempenho, situados no 90º percentil, e aqueles com pior desempenho, pertencentes ao 10º percentil, o que pode ser atribuído, principalmente, ao decréscimo no diferencial do desempenho na cauda superior da distribuição (90-50). Isto significa que parece haver uma queda na “desigualdade” de desempenho entre os estudantes da 4ª série do ensino fundamental, principalmente para aqueles com maiores habilidades cognitivas. Ao decompor esta medida de redução do hiato, observa-se que o efeito retorno e o efeito resíduo foram importantes neste processo.

TABELA 6
Decomposição da variação do desempenho escolar nas medidas de diferença, para escolas e alunos da 4ª série do ensino fundamental avaliados em Matemática
Brasil – 1997-2005

Medidas de diferença	Total	Observáveis		Resíduo	Nível socioeconômico (NSE)	
		Composição	Retorno		Composição	Retorno
Alunos						
1997-1999						
90-10	-5,48	8,74	-8,26	-5,97	6,90	-5,31
90-50	-4,58	4,78	-5,70	-3,66	3,94	-3,77
50-10	-0,90	3,96	-2,55	-2,31	2,96	-1,54
1997-2001						
90-10	6,17	7,34	-1,18	0,01	7,09	-3,81
90-50	2,35	5,17	-2,24	-0,58	4,00	-2,94
50-10	3,83	2,17	1,06	0,59	3,09	-0,87
1997-2003						
90-10	4,28	7,68	-3,34	-0,05	7,87	-1,74
90-50	0,26	4,17	-1,99	-1,92	4,56	-1,29
50-10	4,02	3,51	-1,35	1,87	3,32	-0,45
1997-2005						
90-10	11,71	8,48	-1,54	4,77	7,19	-1,47
90-50	4,72	4,20	-1,47	1,99	3,26	-0,83
50-10	6,99	4,28	-0,07	2,78	3,93	-0,64
Escolas						
1997-1999						
90-10	15,65	35,66	-33,22	13,21	27,72	-33,39
90-50	4,97	17,74	-19,53	6,77	15,77	-22,31
50-10	10,68	17,93	-13,69	6,44	11,95	-11,08
1997-2001						
90-10	25,25	33,41	-12,27	4,10	24,14	-16,88
90-50	10,94	15,20	-6,33	2,06	11,92	-10,72
50-10	14,31	18,21	-5,94	2,04	12,21	-6,16
1997-2003						
90-10	23,28	39,55	-20,77	4,50	32,32	-24,55
90-50	7,47	17,47	-11,48	1,48	15,19	-15,32
50-10	15,81	22,08	-9,28	3,02	17,13	-9,23
1997-2005						
90-10	18,10	27,68	-14,87	5,29	19,33	-17,97
90-50	1,18	6,95	-7,67	1,90	6,59	-9,31
50-10	16,92	20,73	-7,20	3,39	12,73	-8,66

Fonte: Inep. Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

Como o número de matrículas se expandiu em 5,95%¹⁴ entre 1997 e 1999, o suporte da distribuição se alterou, devido à entrada dos alunos com menor desempenho potencial. Neste sentido, houve um inchaço na

cauda inferior da distribuição e, portanto, qualquer quantil em anos posteriores foi diferente (ou pior, em termos do nível de desempenho escolar) quando comparado ao quantil referente a 1997. Este processo

¹⁴ Exclusive alunos de escolas rurais e federais. Dados do Sistema de Avaliação da Educação Básica.

pode explicar, em parte, os resultados encontrados neste biênio.

Nos demais períodos (1997-2001, 1997-2003, 1997-2005), é possível notar um aumento na diferença entre os percentis. Neste cenário, o período de 1997-2005 chama atenção devido ao aumento de 11,7 pontos na diferença entre o desempenho dos alunos do 90º e do 10º percentis. O crescimento na desigualdade entre os percentis se deve, principalmente, ao efeito composição do nível socioeconômico. Isto indica que há uma concentração de alunos mais pobres e com menor desempenho potencial na cauda inferior da distribuição, fato que evidencia a persistência do problema associado à desigualdade de oportunidades educacionais.

Escolas como unidade de análise

Em todos os intervalos analisados, é possível constatar aumento na diferença medida pelo hiato entre o desempenho das escolas situadas no 90º percentil e o desempenho daquelas pertencentes ao 10º percentil da distribuição. Houve, portanto, crescimento na heterogeneidade entre as escolas, no sentido de que ocorreu distanciamento do nível de desempenho alcançado pelas piores e melhores escolas.

Quando analisadas as diferenças abaixo (50-10) e acima (90-50) da mediana, percebe-se que o peso da diferença do desempenho das escolas situadas no segmento inferior da distribuição foi maior. Isto significa que o aumento na heterogeneidade das escolas, em termos do nível médio de aprendizado de seus alunos, foi maior entre aquelas com qualidade mais baixa. Em parte, este efeito pode estar associado ao processo de expansão do ensino, pois o crescimento das matrículas se deu entre os alunos carentes que estudam em escolas de menor qualidade.

A decomposição do hiato mostra que seu aumento pode ser explicado pela soma dos efeitos composição e não-observáveis (resíduo). Nos quatro períodos analisados e nas três medidas de diferença (90-10; 90-50; 50-10), é possível notar que estes efeitos apresentam sinais positivos, contribuindo,

portanto, para elevar a diferença no desempenho escolar entre as escolas.

No que tange ao efeito composição observável (relacionado aos atributos escolares e aos atributos médios dos alunos da escola incluídos na regressão), percebemos novamente que o nível socioeconômico médio das escolas é o componente que mais contribui para a magnitude deste efeito. Em outras palavras, a diferença no nível socioeconômico médio entre as melhores e as piores escolas atuou no sentido de elevar a distância entre elas, em termos de seus resultados acadêmicos.

Já o efeito retorno contribui para reduzir o hiato em todos os períodos. Entre os componentes que o compõem, destaca-se o efeito de mudanças no coeficiente do nível socioeconômico, como pode ser visto na última coluna da Tabela 6.

Se, por um lado, o declínio do coeficiente do nível socioeconômico ao longo dos anos parece mostrar que, de fato, houve perda na qualidade da educação brasileira, por outro, este resultado pode ser interpretado como uma mudança positiva, na medida em que implica reduções nas desigualdades educacionais. Nesse sentido, o declínio da importância das origens sociais sobre os resultados escolares poderia representar um grande avanço se fosse compensado por um aumento da importância das escolas na educação de seus alunos. Isto significa que a escola deveria ser capaz de superar as deficiências e desvantagens trazidas pelos alunos mais pobres, no que tange às suas habilidades cognitivas, melhorando a qualidade média do ensino.

Discussão

Não há resposta simples para a questão de como as mudanças composicionais do público escolar afetaram a média e a distribuição do desempenho escolar dos alunos e das escolas.

Os resultados apresentados neste estudo mostram fortes evidências de que o nível socioeconômico foi o fator mais importante para explicar as variações ocorridas na média e na distribuição dos resultados escolares dos alunos e das escolas nos testes

padronizados de conhecimento entre 1997 e 2005 no Brasil. A influência deste atributo sobre a variação nos resultados escolares depende de qual efeito está sendo analisado: composição ou retorno.

O efeito composição é de entendimento fácil, pois sua influência é direta: quando se aumenta a proporção de estudantes com menor *background* familiar, eleva-se a proporção daqueles com maiores dificuldades de aprendizado e, conseqüentemente, reduz-se o desempenho escolar médio global. Esta dinâmica explica 44% e 87% do declínio da proficiência escolar média entre 1997 e 1999, nas análises micro (aluno) e macro (escola), respectivamente. Nos demais períodos, este componente perde importância relativa. Isto mostra que, de fato, parece haver uma relação direta entre a expansão do acesso ao ensino e a redução da qualidade da educação escolar na segunda metade da década de 1990.

A outra parcela de explicação para o declínio da proficiência média provém do efeito retorno do nível socioeconômico. Este efeito baseia-se nas variações ocorridas nos coeficientes estimados para o nível socioeconômico médio ao longo dos anos e sua interpretação é mais complexa. Os resultados evidenciaram redução da magnitude deste coeficiente nos anos que se seguiram a 1997. Como consequência, houve decréscimo da proficiência média, uma vez que o coeficiente é utilizado como um tradutor para transformar o nível socioeconômico em desempenho escolar. Em outras palavras, o efeito retorno do nível socioeconômico contribuiu de modo expressivo para diminuir a proficiência escolar média nos anos posteriores a 1997. Na análise dos alunos, o efeito retorno do nível socioeconômico explica de 37% a 100% o declínio das médias, dependendo do período analisado. Já na análise das escolas, este efeito é capaz de explicar a totalidade do decréscimo em todos os períodos.

A redução da importância do nível socioeconômico, embora tenha contribuído para diminuir a qualidade média do ensino, atuou no sentido de equalizar a distribuição do desempenho escolar. Isto porque, dado um grau de desigualdade em nível socioe-

conômico, quanto menor a sensibilidade do retorno desta variável, menor é a desigualdade de resultados. De fato, a decomposição das medidas de diferença 90-10, 90-50 e 50-10 mostra que a redução do coeficiente estimado para a média do nível socioeconômico contribuiu para diminuir as diferenças encontradas na proficiência escolar dos alunos e das escolas ao longo dos percentis da distribuição, em todos os períodos analisados. Entretanto, é importante lembrar que o efeito retorno nos percentis foi calculado com base no coeficiente estimado para a média. Esta é uma limitação do presente estudo. Uma análise mais conclusiva do efeito retorno sobre a distribuição requer o uso de outros métodos, como as regressões quantílicas.

Uma hipótese que pode ser levantada para explicar a mudança na sensibilidade do desempenho escolar ao nível socioeconômico relaciona-se ao efeito dos pares, ou seja, às externalidades negativas geradas pela entrada de alunos com menores habilidades e competências cognitivas. Esta externalidade, ocasionada pela interação entre os colegas, tenderia a reduzir o aproveitamento dos alunos com elevado desempenho potencial (por exemplo, alunos com nível socioeconômico mais elevado) e, neste caso, tenderia a igualar os níveis de aprendizado entre os estudantes.

Uma evidência que sustenta a hipótese do efeito de pares negativos foi encontrada em Rodrigues (2009). Ao comparar as distribuições do desempenho escolar observadas em 1997 e 1999, a autora constatou que houve redução na desigualdade (denominada polarização negativa) dos resultados dos alunos avaliados pelo Saeb na 4ª série do ensino fundamental, provocada, principalmente, por uma piora nos resultados educacionais dos alunos situados na cauda superior da distribuição, que, em tese, apresentam melhores condições socioeconômicas e maiores habilidades cognitivas. Ainda com base no Saeb de 2003, Machado et al. (2008) encontraram evidências de que o aumento na heterogeneidade dos alunos afetou negativamente a proficiência escolar individual em Português e Matemática conjuntamente. A

explicação dada pelos autores é de que o aumento na heterogeneidade tende a dificultar a implantação de projetos comuns de aprendizado, dada a diversidade de interesses entre os estudantes.

Em síntese, os resultados apresentados neste estudo parecem reforçar a ideia da existência de um *trade-off* entre a democratização das oportunidades educacionais e a garantia da qualidade da educação escolar.

Referências

ALBERNAZ, A.; FERREIRA, F.H.G.; FRANCO, C. **Qualidade e equidade na educação fundamental brasileira**. Rio de Janeiro: Departamento de Economia da PUC-Rio, maio 2002 (Texto para discussão, 455).

ALVES, F. Qualidade da educação fundamental: integrando desempenho e fluxo escolar. **Ensaio: avaliação de políticas públicas educacionais**. Rio de Janeiro, v. 15, n. 57, p. 525-542, out./dez. 2007.

AZEVEDO, J. P.; FOGUEL, M. N. **Uma decomposição da desigualdade de rendimentos do trabalho no Brasil: 1984-2005**. Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada, dez. 2006 (Texto para discussão, 1247).

BANCO MUNDIAL. **Diversos caminhos para o sucesso educacional**. Boas práticas e desempenho dos alunos na Prova Brasil: identificando casos bem sucedidos na rede municipal de ensino. Brasília: Inep/MEC, 2008. Relatório n. 43590-BR.

BICKEL, R. **Multilevel analysis for applied research: it's just a regression!** New York: The Guilford, 2007.

BOGOTCH, I.; MIRÓN, L.; BIESTA, G. Effective for what; effective for whom? Two questions SESI should not ignore. In.: TOWNSEND, T. (Ed.). **International handbook of school effectiveness and improvement**. Springer Netherlands, 2007, p. 93-110.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Microdados do Sistema Nacional de Avaliação da**

A principal mensagem subjacente é a de que a massificação do ensino não é um processo trivial, havendo preços de qualidade a serem pagos. O desafio é pensar em estratégias que compensem os efeitos negativos da massificação do ensino, principalmente no âmbito das instituições escolares, uma vez que é de suma importância que todas as crianças e jovens estejam inseridos no ambiente escolar.

Educação Básica 1995, 1997, 1999, 2001, 2003 e 2005. Brasília: Diretoria de Tratamento e Disseminação de Informações Educacionais/Inep/MEC, 2006. 6 CD-ROMS.

BRASIL. Ministério da Educação. **Informe dos resultados comparativos do SAEB 1995, 1997 e 1999**. Brasília: Inep, 2000.

BROOKE, N.; SOARES, J. F. S. (Orgs.). **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: UFMG, 2008.

CREEMERS, B. P. M. Educational effectiveness and improvement: the development of the field in Mainland Europe. In.: TOWNSEND, T. (Ed.). **International handbook of school effectiveness and improvement**. Netherlands: Springer Netherlands, 2007, cap. 12, p. 223-244.

FERRÃO, M. E.; FERNANDES, C. A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4ª série. In.: FRANCO, C. (Org.). **Avaliação, ciclos e promoção na educação**. Porto Alegre: ArtMed, 2001, p.155-172.

FLETCHER, P. R. **À procura do ensino eficaz**. Rio de Janeiro: MEC, 1998.

FRANCO, C. et al. O referencial teórico na construção dos questionários contextuais do SAEB 2001. **Estudos em Avaliação Educacional**, São Paulo, n. 28, p. 39-74, jul./dez. 2003.

FRANCO, C. O SAEB – Sistema de Avaliação da Educação Básica: potencialidades, problemas e desafios. **Revista Brasileira de Educação**, Campinas, n. 17, p. 127-132, maio/ago. 2001.

FERNANDES, R.; NATENZON, P. E. A evolução recente do rendimento escolar das crianças brasileiras: uma reavaliação dos dados do SAEB. **Estudos em Avaliação Educacional**, São Paulo, n. 28, p. 3-22, dez. 2003.

GLEWWE, P.; KREMMER, M. Schools, teachers and education outcomes in developing countries. In: HANUSHEK, E.; WELCH, F. (Eds.). **Handbook of the economics of education**. Amsterdam: North-Holland Elsevier, 2006, v. 2, cap. 16, p. 945-1018.

JUHN, C.; MURPHY, K. M.; PIERCE, B. Wage inequality and the rise in returns to skill. **Journal of Political Economy**, Chicago, v. 101, n. 3, p. 410-442, Jun. 1993.

MACHADO, D. C.; GONZAGA, G.; FIRPO, S. P. **A relação entre a proficiência e dispersão de idade na sala de aula**: a influência do nível de qualificação do professor. Rio de Janeiro: Fundação Getúlio Vargas, 2008. Disponível em: <<http://virtualbib.fgv.br/ocs/index.php/sbe/EBE08/paper/view/538/35>>. Acesso em: 04 maio 2009.

MADAUS, G. F.; AIRASIAN, P. W.; KELLAGHAN, T. Insumos escolares, processos e recursos. In: BROOKE, N.; SOARES, J. F. (Orgs.). **Pesquisa em eficácia escolar**: origem e trajetórias. Belo Horizonte: UFMG, 2008, p. 112-141.

NERI, M.; CARVALHO, A. **Seletividade e medidas de qualidade da educação brasileira 1995-2001**. Rio de Janeiro: EPGE; FGV, 2002. (Ensaio econômico, 463).

PICCHETTI, P.; PONCZEK, V.; SOUZA, A. P. Decomposition of the quality of education in Brazil. In: WALTENBERG, F. D. **The quality of education in Brazil**. São Paulo: FGV, 2008 (Final Report).

RAUNDENBUSH, S. W.; BRYK, A. S. **Hierarchical linear models**. 2nd. Newbury Park: Sage Publication, 2002.

RODRIGUES, C. G. **A relação entre a expansão do acesso ao ensino e o desempenho escolar no Brasil**: evidências com base no SAEB para o período de 1997 a 2005. Tese (Doutorado). Centro de

Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, 2009.

SOARES, J. F.; CÉSAR, C. C.; MAMBRINI, J. Determinantes de desempenho dos alunos do ensino básico brasileiro: evidências do SAEB de 1997. In: FRANCO, C. (Org.). **Promoção, ciclos e avaliação educacional**. Porto Alegre: ArtMed Editora, 2001, p. 121-153.

SOARES, J. F. Cognitive achievement inequalities in Brazil: the role of SES, gender, race and region of residence. In: GLOBAL CONFERENCE ON EDUCATION RESEARCH IN DEVELOPING COUNTRIES, 2005, Praga. **Trabalhos apresentados...** Praga, 2005. Disponível em: <www.preal.org/FIE,2005>. Acesso em: 05 mar. 2007.

SOARES, F. V. Some stylized facts of the informal sector in Brazil in the last two decades. In: ENCONTRO NACIONAL DE ECONOMIA, 32, 2004, João Pessoa. **Anais...** João Pessoa: Anpec, 2004.

SOARES, S.; SOUZA, A. **A demografia das escolas**: decomposição da municipalização do ensino básico. Rio de Janeiro: Ipea, 2003 (Texto para discussão, 940).

SOUZA, P. R. **A revolução gerenciada**: educação no Brasil, 1995-2002. São Paulo: Prentice Hall, 2005.

_____. **Avaliação a serviço da qualidade educativa**. São Paulo: reescrevendo a educação, 2006. Disponível em: <www.reescrevendoaeducacao.com.br/2006/pages.php?recid=44-44k>. Acesso em: 05 jul. 2008.

TEDDLIE, C.; STRINGFIELD, S. A history of school effectiveness and improvement research in the USA focusing on the past quarter century. In: TOWNSEND, T. (Ed.). **International handbook of school effectiveness and improvement**. Netherlands: Springer Netherlands, 2007, cap. 8, p. 131-166.

ULYSSEA, G. **Segmentação no mercado de trabalho e desigualdade de rendimentos no Brasil**: uma análise empírica. Rio de Janeiro, 2007 (Texto para discussão, 1261). Disponível em: <<http://www.ipea.gov.br/default.jsp>>. Acesso em: 13 dez. 2009.

ANEXO

Resultados das regressões para escolas e alunos da 4ª série do ensino fundamental avaliados em Matemática Brasil – 1997-2005

Variáveis do aluno e da escola	Mínimos Quadrados Ordinários											
	1997		1999		2001		2003		2005		Erro-padrão	
	Coefficiente	Erro-padrão	Coefficiente	Erro-padrão	Coefficiente	Erro-padrão	Coefficiente	Erro-padrão	Coefficiente	Erro-padrão		
Modelo aluno												
Aluno do sexo feminino	-5,02***	1,28	-1,26	1,09	-4,62***	0,70	-4,01***	0,85	-2,44**	1,12		
Sem resposta no sexo do aluno	-12,92	8,27	-8,37*	5,03	-13,41***	2,17	-21,19***	4,49	-27,12***	3,28		
Aluno autodeclarado como pardo	-1,59	1,48	-1,19	1,24	-2,75***	0,84	-0,89	0,95	-4,71***	1,36		
Aluno autodeclarado como preto	-13,01***	2,11	-16,30***	1,83	-12,80***	1,14	-13,88***	1,41	-17,91***	2,13		
Sem resposta na cor/raça do aluno	-20,21***	4,62	-13,99***	4,76	-18,53***	1,69	-21,34***	3,37	-9,18***	2,55		
Aluno que não reside com ambos os pais	-5,59***	1,50	0,64	1,22	-4,41***	0,77	-0,78	0,84	-0,32	1,06		
Sem resposta na co-residência do aluno	-12,80	9,55	-12,21**	5,20	-15,43***	1,33	-10,36***	3,83	-13,31***	2,16		
Aluno que não faz dever de casa	-15,67***	2,34	-12,09***	1,86	-23,62***	1,38	-14,33***	2,23	-23,01***	2,63		
Sem resposta no dever de casa do aluno	17,53***	5,17	4,02	3,10	4,20**	1,65	2,66	3,53	15,77***	3,14		
Aluno com defasagem idade-série	-19,89***	1,57	-15,05***	1,12	-17,05***	1,03	-17,19***	1,04	-16,63***	1,13		
Sem resposta na defasagem idade-série	-17,65***	2,57	-18,24***	5,43	-22,97***	1,51	-24,14***	5,55	-18,11***	4,02		
Nível socioeconômico do aluno	10,87***	0,90	6,98***	0,43	8,30***	0,33	9,95***	0,41	9,92***	0,48		
Intercepto	157,80***	3,95	163,47***	2,23	161,10***	1,92	145,65***	2,04	152,79***	2,79		
Modelo Escola												
Proporção de alunos do sexo feminino	-8,07	9,72	-4,45	2,98	-4,22	3,57	-4,09	3,65	0,77	4,62		
Proporção de alunos pardos	-10,16	6,90	-2,30	3,04	0,74	2,93	-6,61**	2,76	-1,73	3,99		
Proporção de alunos pretos	-19,92	14,62	-26,18***	5,10	-12,83**	5,74	-11,72**	5,75	-17,03**	7,15		
Proporção de alunos que não residem com ambos os pais	-2,55	9,66	-1,43	2,87	-12,29***	2,89	1,69	3,49	-5,78	5,16		
Proporção de alunos que não fazem dever de casa	-36,22***	11,14	-19,75***	3,17	-45,31***	3,83	-44,98***	6,03	-40,35***	9,48		
Proporção de alunos com defasagem idade-série	-11,66*	6,02	-5,58**	2,85	-6,38**	3,10	-4,65	3,32	-12,04**	5,24		
Nível socioeconômico médio do aluno	17,77***	2,48	6,77***	0,78	11,61***	0,75	10,52***	1,03	9,99***	1,14		
Escola estadual	-5,05	4,01	-17,53***	2,52	-9,94***	2,35	-14,77***	2,38	-16,57***	2,88		

(continua)

(continuação)

Variáveis do aluno e da escola	Mínimos Quadrados Ordinários														
	1997			1999			2001			2003			2005		
	Coefficiente	Erro-padrão		Coefficiente	Erro-padrão		Coefficiente	Erro-padrão		Coefficiente	Erro-padrão		Coefficiente	Erro-padrão	
Escola municipal	-6,58	4,23		-15,31***	2,67		-11,47***	2,09		-17,52***	2,72		-17,92***	3,00	
Infraestrutura escolar	0,60	0,52		2,43***	0,75		0,83***	0,27		0,77***	0,30		-0,08	0,47	
Escola com 50% a 100% dos professores com ensino superior ou mais	2,66	2,58		1,02	1,57		-0,07	1,22		1,03	1,17		2,92*	1,66	
Escola sem a informação sobre a escolaridade dos professores	-2,85	4,48		-2,29	8,66		-0,73	2,15		1,14	4,91		1,43	4,99	
Escola com 60% a 100% de professores com 10 anos ou mais em atividades de docência	1,36	1,98		0,28	1,31		0,42	1,00		1,10	1,18		-1,13	1,72	
Escola sem a informação sobre a experiência profissional dos professores	-	-		0,56	9,09		3,01	4,96		4,25	3,25		-0,85	5,26	
Escola cujo diretor possui curso superior ou mais	-3,15	2,05		-1,26	1,44		-0,11	1,29		-2,30	1,45		-1,65	2,70	
Escola sem a informação sobre a escolaridade do diretor	-2,65	6,49		-1,26	2,63		2,78	3,99		-6,86**	3,30		-6,88*	3,80	
Escola cujo diretor trabalha há 10 anos ou mais em atividades de direção	4,58*	2,36		1,78	1,57		-0,37	1,19		2,17*	1,31		2,70	1,98	
Escola sem a informação sobre a experiência profissional do diretor	0,49	9,02		-6,31*	3,72		-6,00	4,94		6,51	8,59		-6,45*	3,35	
Escola da Região Nordeste	-10,64***	2,64		-8,49***	2,11		-6,64***	1,65		-10,44***	2,07		-16,76***	2,40	
Escola da Região Norte	-14,03***	2,70		-7,78***	2,11		-10,19***	1,86		-12,81***	2,29		-14,06***	2,43	
Escola da Região Centro-Oeste	-3,50	2,65		-2,68	2,19		-3,47**	1,68		-7,17***	1,82		-8,63***	2,49	
Escola da Região Sul	-7,00**	2,84		-4,54**	2,15		-3,89**	1,86		-4,80**	1,89		-3,37	2,68	
Intercepto	137,22***	16,27		169,84***	7,91		153,34***	5,98		157,76***	7,67		174,26***	9,40	
R-ajustado modelo aluno	20,02			18,27			27,83			22,61			20,70		
R-ajustado modelo escola	79,54			49,76			70,72			67,22			64,53		
Amostra de aluno	18.588			16.811			50.782			40.596			37.719		
Amostra de escola	698			2.898			3.551			2.915			2.508		

Fonte: INEP, Sistema de Avaliação da Educação Básica – Saeb, 1997, 1999, 2001, 2003 e 2005.

Nota: Coeficientes: ***significativo a 1%; **significativo a 5%; *significativo a 10%.

Resumen

Diferencias intertemporales en la media y distribución del rendimiento escolar en Brasil: el papel del nivel socioeconómico, 1997 a 2005

Existe un consenso en la literatura de evaluación educacional por el que el nivel socioeconómico de los alumnos es el factor más importante para explicar, en un determinado punto en el tiempo, las variaciones en los resultados escolares de los alumnos evaluados por el Sistema de Evaluación de Educación Básica – Saeb. No se sabe, no obstante, en qué medida el cambio temporal en las condiciones socioeconómicas influyó el cambio en la media y distribución del rendimiento escolar entre los diversos ciclos del Saeb. Para investigar esta cuestión, utilizamos el método de descomposición contrafactual, desarrollado por Juhn, Murphy y Pierce (1993), que permite aislar la contribución de la variación en la composición y retorno del nivel socioeconómico sobre la variación en la media y distribución del rendimiento escolar. Fueron usados los datos del Saeb de 1997 a 2005, para la cohorte de alumnos del 4º curso de enseñanza básica evaluados en matemáticas. Los principales resultados muestran que el aumento en la cobertura escolar al final de los años 1990 redujo el nivel socioeconómico medio de los alumnos y colaboró a disminuir la media y elevar la desigualdad entre los resultados escolares (efecto composición). Al mismo tiempo, hubo una reducción en la sensibilidad del desempeño escolar en el nivel socioeconómico (efecto retorno). Este hallazgo tiene implicaciones ambiguas: por un lado, contribuye a reducir el aprovechamiento escolar medio y, por otro, colabora para equilibrar la distribución de los escolares entre 1997 y 2005.

Palabras-clave: Rendimiento escolar. Nivel socioeconómico. Descomposición contrafactual.

Abstract

Intertemporal differences in averages and distributions of school performance in Brazil: the role of socioeconomic level, 1997-2005

It is general consensus in the literature on the evaluation of education that the socioeconomic level of students is the most important factor for explaining, at any given moment, the variations in the school performance of students evaluated by the Basic Education Evaluation System (SAEB). It is not known, however, how strongly temporal change in socioeconomic conditions affect changes in averages and distributions of school performance among the several different cycles of the SAEB. To investigate this question we used the method of counter-factual decomposition, developed by Juhn, Murphy and Pierce (1993). This method enables researchers isolate the contribution of the variation in the composition and return of socioeconomic level on the variation in the averages and distributions of school performance. The data used was taken from the Saeb applied in 1997 to 2005 to the cohort of students in the fourth grade of elementary education evaluated in mathematics. The main findings show that the increase in school coverage in the late 1990s lowered the average socioeconomic level of the students, was a factor in the reduction of a rise in inequality in school performance (composition effect). At the same time, there were reductions in the sensitivity of school performance to socioeconomic level (return effect). This result has ambiguous implications. On the one hand, it contributes to a reduction in average school performance and, on the other, helps equalize the distribution of scores between 1997 and 2005.

Keywords: School performance. Socioeconomic level. Counterfactual decomposition.

Recebido para publicação em 30/09/2010

Aceito para publicação em 18/02/2011