

# ESFORÇO DA GESTÃO ESCOLAR NA REDE PÚBLICA DE ENSINO FUNDAMENTAL E O DESEMPENHO DOS ESTUDANTES NO BRASIL

Aléssio Tony Cavalcanti de Almeida<sup>1</sup>

Ignácio Tavares de Araújo Júnior<sup>2</sup>

Hilton Martins de Brito Ramalho<sup>3</sup>

Este trabalho investiga os fatores associados ao desempenho dos alunos de escolas públicas brasileiras em exames padronizados, mostrando, em especial, como o esforço empreendido na gestão escolar resulta em ganhos de aprendizagem para o discente. A metodologia consiste na estimação de uma função de produção educacional, a partir dos microdados da Prova Brasil, por meio de um modelo hierárquico linear com três níveis: aluno, escola e município. Os resultados mostram que as diferenças entre os alunos ainda explicam a maior parcela da variabilidade das notas, tanto controlando pelas heterogeneidades entre as escolas quanto pelas diferenças municipais. Entretanto, os rendimentos dos estudantes relacionam-se diretamente com a elevação do desempenho médio da escola nos anos anteriores e com a maior presença de professores efetivos no quadro docente, assim como associa-se inversamente com o histórico de reprovação e/ou abandono na escola. Portanto, o esforço da escola é relevante para explicar o desempenho do aluno, resultado que persiste mesmo após se controlar diferenças entre as municipalidades de cada estado brasileiro.

**Palavras-chave:** qualidade da educação; função de produção escolar; modelo hierárquico.

## EFFORTS OF SCHOOL MANAGEMENT IN THE PUBLIC ELEMENTARY SCHOOL AND STUDENT PERFORMANCE IN BRAZIL

This paper investigates the factors associated with the performance of students in Brazilian public schools on standardized tests, showing, especially, as the efforts made in school management results in learning gains for students. The methodology consists in estimating an educational production function, from the microdata of Prova Brasil, using hierarchical linear model with three levels: student, school and municipality. The results show that the differences between the students explain the largest portion of the variability of scores, both controlling the heterogeneity between schools as by municipal differences. On the other hand, the scores of students are directly related to the increase in the average school performance in previous years and the greater presence of effective teachers in the teaching staff, as well as the historical of repetition and dropout in school has a negative association. Therefore, the school effort is relevant in explaining student achievement, a result that persists even when controlling differences between the municipalities of each Brazilian state.

**Keywords:** quality of education; educational production function; hierarchical model.

---

1. Doutor em economia pelo Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba (PPGE-UFPB). Professor do PPGE-UFPB e pesquisador do Laboratório de Estudos em Microeconomia Aplicada (Lema-UFPB). *E-mail:* <alessio@ccsa.ufpb.br>.

2. Doutor em economia pelo Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco (Pimes/UFPE). Professor do PPGE-UFPB e pesquisador do Lema-UFPB. *E-mail:* <ignacio.tavares@gmail.com>.

3. Doutor em economia pelo Pimes/UFPE. Professor do PPGE-UFPB e pesquisador do Lema-UFPB. *E-mail:* <hilton@ccsa.ufpb.br>.

## LOS ESFUERZOS DE LA GESTIÓN ESCOLAR EN LA ESCUELA PRIMARIA PÚBLICA Y EL DESEMPEÑO ESTUDIANTIL EN BRASIL

Este trabajo investiga los factores asociados con el rendimiento de los estudiantes de escuelas públicas brasileñas en las pruebas estandarizadas, mostrando, en particular, en qué medida los esfuerzos realizados en los resultados de gestión escolar genera mejoras en el aprendizaje para el estudiante. La metodología consiste en la estimación de una función de producción, a partir de los microdatos de la Prova Brasil, cuya estimación se ha realizado mediante un modelo jerárquico lineal con tres niveles: los estudiantes, la escuela y los municipales. Los resultados muestran que las diferencias entre los estudiantes explican la mayor parte de la variabilidad de las puntuaciones, tanto en el control de la heterogeneidad entre las escuelas como por las diferencias municipales. Los resultados muestran que las diferencias entre los estudiantes explican la mayor parte de la variabilidad de las puntuaciones, tanto en el control de la heterogeneidad entre las escuelas como por las diferencias municipales. Por otro lado, las puntuaciones de los estudiantes están directamente relacionados con el aumento en el rendimiento escolar promedio en años anteriores y la mayor presencia de profesores efectivos en el personal docente, así como el histórico de la repetición y de deserción en la escuela tiene una asociación negativa. Por lo tanto, el esfuerzo de la escuela es relevante para explicar el rendimiento de los estudiantes, un resultado que persiste incluso cuando se controlan las diferencias entre los municipios de cada estado brasileño.

**Palabras clave:** calidad de la educación; función de producción educativa; modelo jerárquico.

## LES EFFORTS DE GESTION DE L'ÉCOLE À L'ÉCOLE PRIMAIRE PUBLIQUE ET LE RENDEMENT DES ÉLÈVES AU BRÉSIL

Ce document examine les facteurs associés à la performance des élèves dans les écoles publiques brésiliennes sur des tests standardisés, montrant, en particulier, que les efforts déployés résultent de la gestion de l'école en des gains d'apprentissage des élèves. La méthodologie consiste à estimer une fonction de production de l'éducation, de la microdonnées de Prova Brasil, en utilisant le modèle linéaire hiérarchique à trois niveaux: étudiant, l'école et la municipalité. Les résultats montrent que les différences entre les élèves expliquent la plus grande partie de la variabilité des scores, à la fois le contrôle de l'hétérogénéité entre les écoles que par les différences municipaux. D'autre part, les résultats des élèves sont directement liés à l'augmentation de la performance de l'école moyenne des années précédentes et de la présence accrue des enseignants efficaces dans le personnel enseignant, ainsi que l'historique de redoublement et d'abandon à l'école a une association négative. Par conséquent, l'effort de l'école est pertinent pour expliquer la réussite des élèves, un résultat qui persiste même lorsque le contrôle de différences entre les municipalités de chaque État brésilien.

**Mots-clés:** qualité de l'éducation; fonction de production de l'éducation; modèle hiérarchie.

JEL: I21; H52; H40.

### 1 INTRODUÇÃO

Nos últimos anos, ocorreram notáveis avanços na escolaridade da população brasileira. Segundo dados dos últimos Censos Demográficos de 2000 e de 2010, realizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2000; 2010), a taxa de analfabetismo para a população com idade entre 15 e 24 anos diminuiu de 5,8% para 2,5%. Esse resultado, possivelmente, está associado ao processo

de universalização do ensino fundamental no país. Destarte, mantida a trajetória de aumento do atendimento escolar, muito provavelmente o Brasil atingirá um patamar de escolaridade compatível com grandes economias mundiais e, ao mesmo tempo, poderá reduzir seus graves e persistentes problemas sociais.

A despeito das políticas educacionais voltadas para o aumento da escolaridade média da população brasileira, a qualidade da educação pública no Brasil ainda está abaixo da observada nos países desenvolvidos. Segundo o relatório do Programa Internacional de Avaliação de Alunos (Pisa), em 2009 o Brasil encontrava-se na 53ª posição entre os 65 países avaliados, sendo superado por países latino-americanos como México, Chile e Uruguai (OECD, 2010). Nesse contexto, o aumento da escolaridade média da população pode não ter a efetividade esperada no crescimento econômico, uma vez que há evidências bem documentadas a respeito da relação direta de habilidades cognitivas com retornos salariais, melhor distribuição de renda e maior crescimento econômico (Hanushek e Woessmann, 2008). Solmon (1985), por exemplo, já registrou evidências nessa direção, ao explicar por que, em alguns casos, mais anos de estudo e mais recursos despendidos em educação não necessariamente resultam em maiores ganhos individuais. Hanushek e Kimko (2000), por outro lado, mostram uma associação positiva entre escores de avaliações padronizadas aplicadas em alunos de diversos países e taxas de crescimento econômico, especialmente quando suas estimativas são ponderadas por variáveis que captam diferenças em qualidade de ensino. Desta forma, os autores sugerem que a qualidade da força de trabalho tem uma consistente, estável e forte relação com o crescimento econômico, achado semelhante ao encontrado por Barro (2001). Mais recentemente, Breton (2011) confirma que tanto aspectos quantitativos quanto qualitativos relacionados à educação contribuem para o crescimento econômico.

Ao se cotejar as diferenças inter-regionais dos indicadores de qualidade da educação no Brasil, é possível observar um cenário ainda mais preocupante. Dados do Índice de Desenvolvimento da Educação Básica (Ideb), providos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep, 2009a), por exemplo, para o ano de 2009, revelam que as localidades menos desenvolvidas apresentam, em geral, os piores indicadores quantitativos e qualitativos educacionais. Além disso, ao se considerar apenas os estudantes de escolas públicas na primeira fase do ensino fundamental, as regiões Norte e Nordeste do país registraram médias de 3,7 e 3,5, respectivamente, em uma escala de 0 a 10. Esses valores situam-se muito aquém daqueles observados para as regiões mais desenvolvidas: Sudeste (5,1), Sul (5,0) e Centro-Oeste (4,7). Levando em conta evidências recentes sobre a importância das diferenças regionais de capital humano na explicação das diferenças inter-regionais de renda no Brasil (Barros, 2011), a constatação de tamanha diferença na qualidade da educação entre as regiões Norte

e Nordeste com o restante do país põe dúvida sobre a capacidade dos aumentos de escolaridade nessas duas regiões traduzirem-se nos ganhos de produtividade necessários para diminuir as disparidades inter-regionais de renda *per capita*.

Dado o exposto, torna-se pertinente investigar os determinantes da qualidade de educação no Brasil por meio da análise do desempenho escolar dos alunos brasileiros, verificando, em especial, as seguintes questões de interesse:

- em que medida as diferenças de *performance* individual estão relacionadas com o perfil do aluno e com o *background* socioeconômico familiar?
- qual a parcela da nota na Prova Brasil 2009 do discente imputada à gestão escolar?
- como outras variáveis, que estão sob o controle dos formuladores de políticas educacionais – tais como formação dos professores, situação trabalhista, plano de desenvolvimento da educação, infraestrutura da escola etc. – e que, em última instância, podem capturar a efetividade das aplicações dos recursos públicos investidos em educação, afetam os resultados de aprendizagem dos alunos?

No Brasil, a partir da divulgação dos dados do Sistema de Avaliação da Educação Básica (Saeb) e de alguns sistemas estaduais de avaliação educacional – como o Sistema Mineiro de Avaliação da Educação Pública (Simave)<sup>4</sup> –, trabalhos como o de Fletcher (1998), Franco, Mandarin e Ortigão (2002), Soares (2003), Soares e Candian (2007), Menezes-Filho (2007), Andrade e Laros (2007), Diaz (2007), Franco (2008), Riani e Rios Neto (2008) e Soares e Alves (2013) procuraram conhecer o papel da escola no desempenho do aluno. Em tais estudos, tenta-se capturar o efeito da escola, ao se controlar características das unidades de ensino, tais como existência de laboratórios de informática, conservação da infraestrutura, existência de plano pedagógico etc. Em linhas gerais, os resultados encontrados sugerem que o efeito-escola, quando existe, é pequeno diante dos efeitos na variabilidade das notas ocasionados pelas diferenças socioeconômicas entre os alunos. Com relação aos gastos públicos, os estudos empíricos em destaque não mostram que uma maior alocação de recursos financeiros implica melhor desempenho escolar.

Muito embora se reconheça a valiosa contribuição dos estudos documentados na literatura especializada, ainda há uma lacuna no que diz respeito à identificação de relação causal entre atributos da escola, gasto público em educação e desempenho escolar. Em parte, essa dificuldade pode ser atribuída à existência de variáveis não observáveis, a exemplo do empenho dos responsáveis locais em termos de administração e implementação de boas práticas de ensino direcionadas

---

4. O Simave foi criado no ano 2000 pela Secretaria de Estado da Educação de Minas Gerais.

à diminuição do abandono escolar e a aspectos político-institucionais atrelados ao excesso de professores temporários. Entretanto, os dados contábeis dos municípios disponibilizados pela Secretaria do Tesouro Nacional (Brasil, 2010) não permitem identificar se os recursos financeiros foram efetivamente alocados para aprimorar os meios capazes de resultar em melhorias no desempenho dos alunos.

Este estudo tenta preencher tal lacuna, ao considerar que o esforço da escola em melhorar sua avaliação média é atribuído a fatores não observáveis e, indiretamente, à efetividade do gasto público na escola. Além dessa análise, procura-se avaliar o impacto dos fatores associados ao *background* da escola pública (proficiência média, padrões de repetência e/ou de abandono) sobre o desempenho individual do aluno concludente da primeira fase do ensino fundamental, levando em conta as diferenças de atributos observados entre os discentes e entre as escolas.

Além desta introdução, este artigo está organizado em quatro seções. A seção 2 apresenta uma revisão da literatura especializada. Na seção 3, são discutidos os aspectos metodológicos deste estudo. Na seção 4, são analisados os principais resultados econométricos. Por fim, a seção 5 é reservada às considerações deste trabalho.

## 2 UMA VISÃO GERAL DA LITERATURA EM ECONOMIA DA EDUCAÇÃO

O estudo pioneiro na estimação da função de produção escolar, também conhecido como *Relatório Coleman*, foi publicado em 1966 (Coleman *et al.*, 1966). Nesse trabalho, foi constatado que, naquela época, as diferenças de desempenho dos alunos na rede pública dos Estados Unidos deviam-se muito mais às diversidades de atributos pessoais dos discentes do que às diferenças entre as unidades de ensino. Ou seja, os professores possuíam um pequeno papel no aprendizado do aluno. Tais resultados foram considerados controversos, e houve críticas à abordagem metodológica adotada, que não levou em conta a existência de variáveis não observáveis, como a motivação do professor e a didática em sala de aula, tornando as estimativas enviesadas e de difícil interpretação. Em estudos mais recentes, Rivkin, Hanushek e Kain (2005) e Clotfelter, Ladd e Vigdor (2007) confirmam a influência do professor e de características da escola no desempenho do aluno.

Com relação aos resultados obtidos na estimação da função de produção escolar para o Brasil, constatou-se que uma parcela importante das diferenças entre as notas dos estudantes nas avaliações padronizadas deve-se às próprias diferenças entre os alunos, resultado que corrobora aqueles registrados na literatura internacional. O efeito da escola existe, porém tem um poder explicativo menor do que as características do aluno. Tais evidências, por exemplo, foram encontradas por Fletcher (1998), utilizando dados da prova de matemática do Saeb de 1995 para alunos da 8ª série, e por Franco, Mandarino e Ortigão (2002), com base nas notas dos alunos da 8ª série no Saeb de 1999, que mostraram que a existência de

projetos pedagógicos de escola não promoveu eficácia escolar e aumentou o efeito da origem social nas notas dos alunos nos testes.

O trabalho de Menezes-Filho (2007) investiga os determinantes do desempenho escolar nas escolas públicas e privadas no Brasil, utilizando como indicador de desempenho as notas na proficiência de matemática do Saeb de 2005. Segundo o autor, mesmo controlando para características socioeconômicas, os alunos das escolas privadas têm um desempenho melhor do que os estudantes das escolas públicas. Além disso, os dados mostram que entre 10% e 30% das diferenças de notas obtidas pelos alunos da rede pública ocorrem devido a diferenças entre escolas. O restante da variação ocorre dentro das escolas, ou seja, devido a diferenças entre os alunos e suas famílias. Com efeito, Menezes-Filho (2007) ressalta que as variáveis que mais explicam o desempenho escolar são o nível educacional da mãe, a cor da pele, o atraso escolar, o número de livros, a presença de computador em casa, o trabalho infantil e a realização de pré-escola. Com relação à escola, o número de computadores, o processo de seleção do diretor e de estudantes, a escolaridade, a idade e o salário dos professores têm efeitos muito reduzidos sobre o desempenho escolar.

Franco (2008) construiu um painel de dados contendo diferentes características dos alunos, dos professores e das escolas, além de registrar as notas médias obtidas pelos alunos da 4ª série em matemática nos anos de 1997, 1999, 2001, 2003 e 2005. Com tal base de dados, a autora mostrou que as características do aluno e de suas famílias explicam mais o desempenho dos estudantes nas escolas particulares do que nas públicas. O efeito negativo do atraso escolar também é mais forte nas escolas privadas. A autora não encontrou efeitos de características observáveis dos professores, tanto para escolas privadas quanto para públicas. Nas escolas públicas, as características do diretor e da turma não explicam o desempenho médio dos alunos nas escolas.

Riani e Rios-Neto (2008) utilizaram um modelo hierárquico de resposta binária na tentativa de conhecer que aspectos da família e da estrutura escolar dos municípios brasileiros contribuem para a probabilidade de o aluno estar na série correta para a sua idade. Conforme os autores, os resultados da análise corroboram a hipótese de que a melhoria da rede escolar dos municípios é um importante fator de diminuição da estratificação educacional, seja pelo seu efeito direto no aumento médio das probabilidades avaliadas, seja por diminuir a importância dos fatores relacionados ao ambiente familiar do aluno, reduzindo a desigualdade intergeracional.

Uma questão importante na avaliação do desempenho dos alunos diz respeito ao impacto do gasto público na qualidade do ensino ou no desempenho escolar deles. Cooray (2009), a partir dos resultados obtidos com dados de 46 países de renda baixa e média, argumenta que novos gastos em educação são importantes, pois eles, além de melhorarem a qualidade do ensino, contribuem para o crescimento econômico.

Porém, tanto na literatura internacional quanto na nacional, não há evidências convincentes de que o aumento do gasto com educação melhora o desempenho do aluno. Hanushek (1995), por exemplo, analisando 96 trabalhos empíricos feitos para países em desenvolvimento, constatou que, entre um total de doze casos investigados, apenas em seis houve efeito positivo do gasto por aluno no desempenho escolar. Avaliando os artigos publicados após 1995 a respeito do tema em destaque, Hanushek (2010) também mostra que poucos estudos encontraram evidências sobre a existência de uma relação causal entre gasto público e resultado escolar. Segundo o autor, essa constatação deixa claro que o uso de recursos nas escolas está sujeito à considerável ineficiência, pois as escolas, sistematicamente, gastam com insumos que não estão correlacionados com desempenho escolar.

Para o caso brasileiro, Diaz (2007) avaliou os possíveis efeitos do gasto municipal em educação no Ideb dos municípios, levando em conta as características dos alunos e da escola e a natureza hierárquica dos dados. Nesse trabalho, constatou-se que a participação dos gastos com educação nas despesas totais não possui um impacto positivo no Ideb. Em outras palavras, a autora destaca que o volume de recursos financeiros destinados para a educação reflete opções políticas que não necessariamente são acompanhadas por mudanças em aspectos que diretamente afetariam as condições ou os fatores que efetivamente são responsáveis pela melhoria da qualidade do ensino público municipal. Nascimento (2007), utilizando as notas dos alunos em provas de matemática e português aplicadas nas escolas baianas em 2000, também constatou que a participação relativa do gasto público em educação não tem efeito sobre o desempenho do aluno. Em suma, os trabalhos comentados sugerem que, no Brasil, a escola não é eficaz e não há garantias de que um aumento dos recursos destinados à educação resulte em tal eficácia.

Recentemente, estudos como os de Soares e Alves (2013), Almeida (2014) e Costa, Arraes e Guimarães (2015) buscaram estudar como a gestão da educação, seja dentro da escola, seja no município, afeta o aprendizado nas escolas brasileiras. Em Soares e Alves (2013), os efeitos da escola e do município no desempenho escolar no ensino fundamental nas escolas públicas do Brasil são captados por meio dos resíduos obtidos na estimação de modelos hierárquicos, controlando-se para os fatores fora do contexto da escola. De forma geral, os autores identificaram que municípios têm papel importante no resultado dos alunos. Conforme indica o estudo, por meio da análise dos resíduos das regressões estimadas encontram-se escolas e municípios que cumprem bem seu papel e agregam conhecimento aos alunos, mas sem apontar como escolas e municípios chegam a tal resultado. Da mesma forma, o estudo também permite identificar escolas que apresentam bons resultados apenas por contar com alunos bem preparados que teriam o mesmo desempenho onde quer que estudassem. Por fim, o estudo apontou a

existência de escolas que conseguem obter níveis adequados de proficiência de alunos dos quais não se esperava obter tal resultado.

Ao estudar como alguns insumos escolares, representados pelos atributos físicos, pedagógicos e de pessoal das escolas, contribuem para o aprendizado escolar, Almeida (2014) conclui que a existência de programas de reforço de aprendizagem para os alunos não influencia a direção esperada da nota dos estudantes. Segundo a argumentação do autor, o simples fato de o diretor da escola apontar que este tipo de programa está sendo desenvolvido não gera as mudanças desejáveis no desempenho escolar dos alunos com as menores notas. Costa, Arraes e Guimarães (2015) estudaram o impacto de uma característica importante da gestão municipal da educação, que é a presença de professores sem estabilidade no sistema escolar local. O estudo mostra que entre os municípios brasileiros, quanto maior a falta de professores com estabilidade, menor é o desempenho dos alunos. De acordo com os autores, tal resultado pode estar relacionado ao mercado de trabalho dos professores que, supostamente, pode ser menos ou mais competitivo, bem como indicar que o benefício não pecuniário da estabilidade é um importante benefício na seleção de profissionais que influenciam a melhoria da qualidade do ensino

Outro fenômeno relacionado à gestão local da educação é o impacto do processo de descentralização na oferta do ensino fundamental recentemente ocorrido no Brasil. Teoricamente, a gestão descentralizada (ver Oates, 1977; 2005) pode resultar em uma série de vantagens, porque possibilita aumentos de eficiência dos governos locais na provisão de produtos e serviços mais próximos das preferências e demandas de suas localidades. Espera-se, portanto, que escolas gerenciadas pelos municípios tenham melhor desempenho escolar do que aquelas sob a gestão da administração estadual. Porém, segundo Ceneviva (2011), mesmo controlando para as características dos alunos, o efeito da municipalização das matrículas sobre o desempenho escolar no Brasil foi quase nulo quando comparado ao resultado obtido pelas escolas sob a gestão do governo estadual. Nessa mesma direção, Leme, Paredes e Souza (2009) também não encontraram resultados que apontam diferenças entre escolas que foram municipalizadas e escolas do ensino fundamental sob a gestão estadual.

Mesmo observando certo rigor econométrico na maioria dos trabalhos reportados, as estimativas obtidas ainda podem ter algum viés pela omissão de insumos educacionais relevantes para a explicação do desempenho dos alunos. Conforme argumentam Clotfelter, Ladd e Vigdor (2007) e Hanushek (2007), uma forma de contornar o viés de omissão de variáveis explicativas relevantes na estimação da função de produção escolar é incluir a variável de desempenho escolar defasada como explicativa no modelo econométrico. Os autores argumentam que o desempenho defasado do aluno seria uma *proxy* para efeitos cumulativos e não

diretamente observados sobre o ambiente escolar, haja vista que o desempenho em anos passados pode estar correlacionado com insumos escolares previamente disponibilizados. Nos trabalhos aplicados para o Brasil, citados neste artigo, não foi observado tal procedimento, havendo sempre a tentativa de correlacionar a nota do aluno no ano  $t$  com os insumos escolares observados no mesmo ano. Em parte, isso se deve ao fato de não haver dados longitudinais a respeito do desempenho dos alunos. Uma alternativa seria avaliar o desempenho médio da escola, cujos resultados passados podem ser verificados nas bases de dados do Ministério da Educação (MEC).

Addonizio (2009) formaliza a ideia de avaliar o desempenho médio da escola em função do desempenho passado, ao considerar uma comutatividade do aprendizado, ou seja, o estoque de conhecimento dependendo dos insumos escolares disponíveis em períodos anteriores. Neste sentido, a função de produção escolar  $f$  pode ser apresentada da seguinte forma:

$$A_{it} = f(S_{it}, S_{it-1}, S_{it-2}, \dots) + \epsilon_{it}, \quad (1)$$

onde  $A_{it}$  representa o desempenho médio da escola no período  $t$ ;  $\epsilon_{it}$  é um termo de erro randômico;  $S_{it}$  representa os recursos educacionais da escola  $i$  no período  $t$ ;  $S_{it-1}$  são recursos da escola no período  $t-1$ ;  $S_{it-2}$  são recursos escolares no período  $t-2$ , e assim por diante.

Assumindo-se que os efeitos dos insumos escolares sobre o resultado do aluno são lineares e constantes entre os anos, e sendo  $\beta$  o efeito dos recursos escolares no desempenho do aluno e  $\alpha$  ( $0 < \alpha < 1$ ) uma constante, que capta a persistência da taxa de conhecimento de um ano para o outro. Ou seja, a equação (1) pode ser apresentada como:

$$A_{it} = \beta S_{it} + \alpha(\beta S_{it-1} + \alpha\beta S_{it-2} + \alpha^2\beta S_{it-3} \dots) + \epsilon_{it}. \quad (2)$$

O termo entre parênteses é, justamente,  $A_{it-1}$ ; portanto, a equação (2) pode ser escrita da seguinte forma:

$$A_{it} = \beta S_{it} + \alpha A_{it-1} + \epsilon_{it}. \quad (3)$$

Conforme a equação (3), o efeito dos recursos defasados no desempenho médio da escola é capturado pelo parâmetro  $\alpha$ . Acrescentando um vetor de características dos alunos na equação (3), Addonizio (2009) constatou que o efeito do desempenho passado da escola é positivo e estatisticamente significativo, a partir de dados de unidades escolares do ensino fundamental do estado de Minnesota,

nos Estados Unidos. Resultados na mesma direção foram encontrados por Clotfelter, Ladd e Vigdor (2007), avaliando um painel de dados de alunos de escolas americanas.

O respaldo teórico-empírico e a perspectiva de contornar o problema de omissão de variáveis motivam a realização de estudos utilizando informações anteriores sobre os alunos ou da escola na função de produção escolar. Este trabalho tenta contribuir neste sentido, admitindo que as mudanças de desempenho escolar no passado, o uso de insumos empregados no passado e a melhoria promovida neles têm papel importante no nível de aprendizado corrente do aluno.

### 3 ESTRATÉGIA EMPÍRICA

Esta seção descreve a metodologia do trabalho, com destaque para as principais características de modelos econométricos multinível (hierárquicos) e descrição da base de dados e dos respectivos tratamentos realizados para obtenção das variáveis incorporadas ao modelo empírico.

#### 3.1 Modelo empírico

A natureza hierárquica de dados sobre educação é bastante comum e pode ser mais bem entendida ao se considerar, por exemplo, que os estudantes a serem analisados estão distribuídos entre turmas que, por sua vez, encontram-se em escolas situadas em um dado município. Deste modo, ao se ignorar o conjunto de informações inerentes à referida estrutura dos dados e à heterogeneidade não observada em diferentes níveis de hierarquia (efeitos contextuais), corre-se o risco de mascarar possíveis relações e de se obter estimadores ineficientes (Bafumi e Gelman, 2006).

Nesse cenário, a análise dos determinantes da *performance* do aluno deveria levar em consideração também o questionamento a seguir: a própria natureza hierárquica das informações educacionais consegue explicar ou não, em alguma medida, certas diferenças no desempenho escolar dos discentes? Em Rasbash *et al.* (2009), pode-se encontrar o seguinte argumento que, em certo grau, ajuda a responder tal indagação:

quando se mede a realização educacional de um aluno, deve-se atentar que, em média, esse desempenho varia de uma escola para outra. A nota obtida por estudantes dentro de uma escola será mais igual, em média, do que a nota de discentes de escolas distintas (Rasbash *et al.*, 2009, p. 2, tradução dos autores).

Destarte, de modo semelhante a diversos estudos na literatura especializada – como Natis (2001), Soares (2003), Franco, Mandarino e Ortigão (2002), Diaz (2007), Riani e Rios-Neto (2008) e Soares e Alves (2013) –, este trabalho adotou um modelo multinível (hierárquico) para tentar captar o efeito da gestão escolar sobre o desempenho individual de cada estudante conluente da

primeira fase do ensino fundamental da rede pública, considerando heterogeneidades observada e não observada em até três grupos: alunos (primeiro nível), escolas (segundo nível) e municípios (terceiro nível). Esse tipo de abordagem realiza regressões em que os dados são estruturados em grupos e os coeficientes podem variar de acordo com cada *cluster* definido.

Segundo Gelman e Hill (2007), os modelos de regressão multinível podem ser usados para uma variedade de propósitos, incluindo inferência causal, previsão e modelagem descritiva. Conforme os mesmos autores, algumas motivações para o uso dessa técnica são realçadas a seguir: *i*) entender como os efeitos de tratamento variam entre os indivíduos de grupos diferentes; *ii*) analisar dados estruturados, pois em alguns conjuntos de informações os próprios dados possuem uma estrutura inerente a vários níveis; por exemplo, os alunos dentro das escolas, os pacientes dentro de hospitais ou dados de amostragem por conglomerados; *iii*) obter inferências mais eficientes para os parâmetros de regressão, dado que nas regressões tradicionais tais diferenças, em especial, são ignoradas entre grandes grupos.

Todavia, cabe ressaltar que a adoção desse tipo de modelagem é justificável apenas se o modelo estatístico reconhecer uma estrutura hierárquica presente nos dados. Caso contrário, as estimações podem ser realizadas em um único nível. Desta maneira, a especificação do modelo é construída a partir de um processo heurístico, em que se procura capturar o grau de variabilidade dos dados entre os grupos hierárquicos.

Para exemplificar esse processo, considere-se, inicialmente, um modelo em nível-aluno para uma dada escola a partir da seguinte equação:

$$y_i = \beta_0 + \beta_1 x_i + e_i, \quad (4)$$

onde  $y_i$  é a nota do teste de matemática;  $x_i$  é uma dada variável explanatória do  $i$ -ésimo aluno;  $e_i$  é o termo de erro, todos indexados para o  $i$ -ésimo aluno;  $\beta_0$  é o intercepto; e  $\beta_1$  é um coeficiente fixo associado ao atributo  $x_i$ .

Por sua vez, uma equação em nível-escola para diversas escolas é dada por:

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij}, \quad (5)$$

onde  $j$  indexa as escolas;  $\beta_{0j}$  representa um intercepto aleatório entre as escolas, isto é,  $\beta_{0j} \equiv \beta_0 + u_{0j}$ , de modo que  $\beta_0$  é uma constante e  $u_{0j}$  é um termo estocástico que mensura o desvio do intercepto da  $j$ -ésima escola. Além disso, assume-se que os termos aleatórios  $u_{0j}$  e  $e_{ij}$  são normalmente distribuídos  $u_{0j} \sim N(0, \sigma_{u0}^2)$  e  $e_{ij} \sim N(0, \sigma_e^2)$ , com respectivas variâncias  $\sigma_{u0}^2$  e  $\sigma_e^2$ .

Em uma análise de dois níveis (alunos e escolas), como a expressa na equação (5), as escolas são tratadas como uma amostra aleatória de uma população de escolas.<sup>5</sup> A presença de dois termos aleatórios  $u_j$  e  $e_{ij}$  na última equação, caracteriza o modelo multinível. Em particular, excluindo-se  $x_{ij}$  da equação (5), chega-se ao chamado modelo nulo. A semelhança entre indivíduos do mesmo grupo pode ser medida pela correlação intraclasse (*variance partition coefficient* – VPC), que mensura a proporção da variação residual total que é explicada por diferenças entre os grupos, ou seja, captura a importância da estrutura hierárquica na explicação da variabilidade total.<sup>6</sup>

Na abordagem multinível, a definição do modelo econométrico tem início estimando a equação (5) na sua forma mais simples possível, isto é, o modelo nulo, para, posteriormente, acrescentar mais variáveis a este modelo. Esse procedimento heurístico permite a seleção da melhor especificação por um ou mais critérios preestabelecidos. Em geral, os critérios mais usados para a construção de um modelo multinível são: estatística Deviance,<sup>7</sup> critérios de informações Akaike (AIC) e Bayesiano (BIC).

Neste artigo, adotou-se o cálculo do VPC como uma forma inicial de verificar a relevância ou não de se considerar os níveis da escola e do município. Além disso, baseando-se em Soares (2003), este estudo levou em conta as estatísticas Deviance e AIC para a seleção do modelo com melhor ajustamento aos dados.

Segundo Rasbash *et al.* (2009), um modelo multinível simples, em que os únicos parâmetros aleatórios são as variâncias do intercepto de cada nível, é conhecido como um modelo de componentes de variância. A especificação de um modelo mais geral que aquele apresentado na equação (5) requer a inclusão de uma variável explicativa especial  $x_0$ , que assume o valor 1 para todos os alunos, e a existência de uma variável  $z_j$  relacionada ao nível superior ao do aluno. Assim, a fim de generalizar (5), a expressão (6), a seguir, apresenta a notação matricial para o modelo multinível:

$$y_j = X_j\beta + Z_ju_j + e_j, \quad (6)$$

onde  $j = 1, \dots, J$  é número de *clusters* independentes;  $y_j$  representa o vetor da variável dependente (nota de matemática dos estudantes) para o *cluster*  $j$ , que corresponde

5. O número de grupos deve ser razoavelmente grande. Se, por exemplo, o número de escolas for pequeno, os efeitos dos grupos poderiam ser capturados incluindo variáveis *dummies* para cada escola em uma equação não hierárquica.

6. Formalmente,  $VPC = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2}$ .

7. A estatística Deviance é dada por  $-2\log(L^*)$ , onde  $\log(L^*)$  é o valor máximo (em logaritmo) da função de verossimilhança. Essa estatística é bastante usada na literatura, pois mede o grau de ajustamento dos dados ao modelo construído (Soares, 2003).

às linhas do vetor  $y$  de ordem  $n \times 1$ ;  $X_j$  é a matriz de variáveis explanatórias  $n_j \times p$  para os efeitos fixos de  $\beta$ ;  $Z_j$  é a matriz de covariáveis  $n_j \times q$  para os efeitos aleatórios  $u_j$  dos níveis superiores ao nível-aluno, com média zero e matriz de variância  $G$  de ordem  $q \times q$ ;  $e_j$  é vetor de erros  $n_j \times 1$  relacionados ao nível-aluno, com média zero e matriz de variância  $R = \sigma_e^2 I_j$ .

É válido ressaltar que a parte  $X\beta$  da equação (6) refere-se ao componente do efeito fixo do modelo, e a soma de  $Zu$  e  $e$  na mesma expressão indica o componente estocástico do modelo atribuído ao primeiro e aos demais níveis considerados. Assume-se que o termo de erro  $u$  tem matriz de variância-covariância  $G$  e que  $u$  é ortogonal para  $e$ . As matrizes  $G$  e  $R$  podem ser modeladas de diferentes formas. Por exemplo, a matriz  $G$  pode ser considerada como independente, ou correlacionada, ou independente com variâncias iguais. No caso deste artigo, o modelo gerado não impôs restrições nem para a matriz  $G$  e nem para a matriz.

Considerando-se o termo estocástico conjunto  $Z_j u_j + e_j$ ,  $y_j$  segue uma distribuição normal multivariada com média  $X_j \beta$  e matriz de covariância  $V = ZGZ^T + R$ . Além disso, seja  $\theta$  um vetor de elementos únicos de  $G$ , então o modelo (6) pode ser estimado a partir da maximização da seguinte função de *log-verossimilhança*:

$$L(\beta, \theta, \sigma_e^2) = -\frac{1}{2} \left\{ n \log(2\pi) + \log|V| + (y_j - X_j \beta)^T V^{-1} (y_j - X_j \beta) \right\}. \quad (7)$$

A otimização eficiente de (7) pode ser obtida subdividindo os dados em grupos e subgrupos independentes, aplicando-se uma decomposição matricial triangular (decomposição de Cholesky) para a matriz de variância de  $u$  e procedimentos iterativos sob a técnica de *expectation-maximization* (EM). Para mais detalhes, ver Pinheiro e Bates (2000).

### 3.2 Base de dados

As informações usadas neste trabalho são oriundas dos microdados da Prova Brasil 2007 e 2009, disponibilizados no sítio eletrônico do Inep (2007; 2009b). A Prova Brasil registra diversas informações sobre características de alunos concluintes do 5º e 9º anos do ensino fundamental, professores, escolas e diretores, permitindo, portanto, a investigação de uma ampla gama de possíveis fatores relacionados ao desempenho escolar.<sup>8</sup>

Em conformidade com o objetivo deste estudo, tornou-se pertinente efetuar uma série de recortes na amostra principal da Prova Brasil 2009. Do total de

8. Com as mudanças estabelecidas na Lei nº 11.274/2006, que alterou a duração do ensino fundamental de oito para nove anos e tornou obrigatória a matrícula para crianças a partir dos 6 anos de idade, tem-se, por exemplo, que a antiga 4ª série é atualmente designada como 5º ano, isto é, o último período da primeira fase do ensino fundamental.

aproximadamente 5,9 milhões de alunos entrevistados, foram considerados apenas aqueles concluintes da primeira fase de ensino fundamental (5º ano), que estudavam em escolas públicas de zonas urbanas e que responderam aos quesitos referentes ao gênero, à raça, ao trabalho, à idade, à reprovação ou ao abandono. Após se fazer o cruzamento dos últimos dados com as informações dos exames de português e de matemática, a amostra resultante foi de 1.878.602 estudantes com notas registradas.

As informações das escolas foram coletadas a partir dos questionários referentes aos alunos, aos professores e às notas dos exames. Considerando apenas as escolas públicas situadas em áreas urbanas com oferta da primeira fase do ensino fundamental, ou o equivalente a 35.339 unidades segundo os dados da Prova Brasil de 2009, contabilizaram-se, entre outros indicadores, as notas médias dos exames, as taxas de reprovação e/ou abandono e as taxas de professores em regime de trabalho estável por escola. Além disso, ao se efetuar o cruzamento dos últimos dados com as notas médias por escola registradas na Prova Brasil de 2007, a amostra reduziu-se para 34.841 escolas, permitindo o cálculo da variação percentual das notas das unidades escolares no período de 2007 a 2009.

O último expediente para a formatação final da amostra consistiu na ligação dos dados dos alunos concluintes do 5º ano com aqueles selecionados para as escolas. Por seu turno, foram considerados apenas os alunos corresidentes com pai e mãe (68% do total de alunos) e excluídos da amostra aqueles que não souberam informar a escolaridade dos seus responsáveis, filtro que possibilitou o cálculo da escolaridade média dos pais. Por fim, excluíram-se os dados referentes a escolas com menos de dez alunos, totalizando uma amostra final de 481.725 alunos, agrupados em 20.545 escolas e em 4.311 municípios.

O quadro 1 apresenta a definição e a descrição do conjunto de variáveis selecionadas para análise empírica considerando as seguintes hierarquias: alunos agrupados em escolas e escolas agrupadas em municípios. Vale destacar que a variável dependente de interesse neste estudo é a nota de matemática do aluno (em logaritmo), pois se pressupõe que a aprendizagem do discente em tal disciplina é influenciada, em especial, pelo papel da escola.<sup>9</sup> No apêndice A são apresentadas as estatísticas descritivas para as variáveis presentes no quadro 1 referentes aos níveis aluno e escola. Vale ressaltar que o termo pré-selecionado empregado no quadro 1 foi usado em virtude do processo heurístico de especificação final do modelo econométrico, em que as variáveis são efetivamente por critérios parcimoniosos de ajustamento.

---

9. Na literatura nacional sobre estudos relacionados aos determinantes do desempenho escolar, encontra-se uma série de trabalhos – como Fletcher (1998), Franco, Mandarino e Ortigã. (2002), Menezes-Filho (2007) e Franco (2008) – que usaram a nota de matemática como variável dependente.

**QUADRO 1**  
**Descrição das variáveis pré-selecionadas para construção do modelo**

Nível	Variável	Descrição
(I) Aluno	<i>ID_ALUNO</i>	Código de identificação do aluno.
	<i>LNOTAM</i>	Logaritmo natural da proficiência no exame de matemática de 2009.
	<i>SEXO</i>	1 = menino. 0 = menina.
	<i>REPROVACAO</i>	1 = aluno já reprovou e/ou abandonou os estudos. 0 = caso contrário.
	<i>DFIDADE</i>	Distorção idade-série: a idade teórica do aluno na 4ª série (ou 5ª ano) seria 10 anos. Assim, para idade > 10, tem-se DFIDADE > 0.
	<i>TRABALHA</i>	1 = aluno trabalha. 0 = caso contrário.
	<i>RACA</i>	1 = aluno que se autodeclarou branco 0 = caso contrário.
	<i>PAIS_ESTUDO</i>	Faixa de escolaridade média dos pais ou responsáveis.
(II) Escola	<i>ID_ESCOLA</i>	Código de identificação da escola.
	<i>T_SITTRAB</i>	Proporção de professores na escola que possui situação trabalhista estável.
	<i>ESC_PEDAPORT</i>	Indicador médio das práticas pedagógicas dos docentes para o ensino de matemática.
	<i>ESC_PROFESSORES</i>	Escolaridade média dos professores.
	<i>ESC_SALESCOLA</i>	Salário médio dos professores na escola.
	<i>ENOTAM</i>	Proficiência média da escola em matemática em 2007.
	<i>DELTA_ENOTAM</i>	Evolução da proficiência média da escola em matemática entre 2007 e 2009.
	<i>SNOTAM</i>	Desvio-padrão da proficiência em matemática da escola entre os seus alunos.
	<i>TDI</i>	Taxa de distorção idade-série para o 5º ano do ensino fundamental.
	<i>T_REPROVACAO</i>	Proporção dos alunos da escola que já foram reprovados ou abandonaram os estudos.
	<i>T_TRABALHA</i>	Proporção dos alunos da escola que trabalham.
	<i>ESC_PAIS_ESTUDO</i>	Escolaridade média dos pais dos alunos de uma dada unidade escolar.
(III) Município	<i>COD_MUNICIPIO</i>	Código de identificação do município.
	<i>IFDM</i>	Índice Firjan de Desenvolvimento Municipal (IFDM) de 2007.
	<i>DELTA_IFDM_1</i>	Variação do IFDM entre 2007 e 2009.
	<i>DELTA_IFDM_2</i>	Variação do IFDM entre 2005 e 2009.

Elaboração dos autores.

## 4 RESULTADOS

### 4.1 Exploração estatística dos efeitos de estruturas hierárquicas

Uma forma de se investigar a necessidade de estimação de um modelo econométrico multinível é decompor a variância em duas parcelas: *i*) a parcela explicada pelas diferenças entre alunos; e *ii*) a parcela imputada às diferenças entre escolas. Tal decomposição é realizada estimando-se um modelo de efeitos aleatórios,

também chamado de modelo nulo, na forma da equação (8). O modelo a seguir representa um modelo nulo de dois níveis: o primeiro inclui os estudantes ( $i = 1, \dots, I$ ) e o segundo as escolas ( $j = 1, \dots, J$ ). A variável dependente do modelo é a nota de matemática do discente. Desta maneira, o resíduo é particionado dentro do componente do nível I ( $e_{0ij}$ ) e do componente do nível II ( $u_{0j}$ ). Neste tipo de modelo, a variância entre os grupos é  $\sigma_{u0}^2$  e a variância entre os indivíduos é  $\sigma_e^2$ .

$$\begin{aligned} y_{ij} &= \beta_{0ij}x_{0ij} \\ \beta_{0ij} &= \beta_0 + u_{0j} + e_{0ij} \\ u_{0j} &\sim N(0, \sigma_{u0}^2) \\ e_{0ij} &\sim N(0, \sigma_{e0}^2), \end{aligned} \tag{8}$$

onde  $y_{ij}$  é a nota individual em matemática (em logaritmo) do aluno  $i$  na escola  $j$ ; e  $x_{0ij}$  é uma constante igual a uma unidade para todos os alunos da amostra.

A semelhança entre indivíduos do mesmo grupo na abordagem de modelos multinível, como salientado na seção referente ao modelo empírico, é medida pelo VPC. Conforme ressaltam Rasbash *et al.* (2009), o número de grupos deve ser razoavelmente grande. Por exemplo, se o número de escolas ( $J$ ) é pequeno, os efeitos dos grupos podem ser capturados incluindo variáveis *dummies* para cada escola.

Para o caso desta pesquisa, que leva em consideração mais de 20 mil escolas, já se tem um indicativo da necessidade da adoção da abordagem hierárquica. Contudo, fazendo o cálculo do VPC, inicialmente, para o caso de dois níveis (aluno e escola) e o conjunto de unidades de ensino no território nacional, tem-se um argumento estatístico que mostra a relevância de se considerar uma estrutura hierárquica para a análise dos determinantes do desempenho escolar.

Ao levar em conta todos os alunos, cerca de 482 mil estudantes, e todas as unidades escolares da amostra selecionada, em torno de 20 mil escolas, não se pode negar a importância das diferenças nas escolas sobre o desempenho em matemática dos discentes, pois a partir da estimação do modelo nulo com dados do Brasil, encontrou-se que quase um quarto da variância total da nota em matemática do aluno pode ser atribuído às diferenças entre unidades públicas de ensino, conforme será discutido mais adiante.<sup>10</sup>

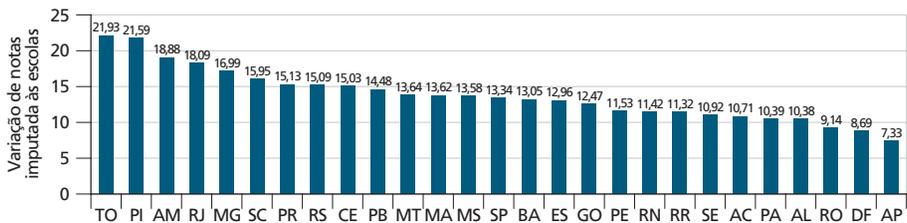
Na tentativa de verificar algum padrão acerca da explicação do efeito-escola sobre a variabilidade das notas dos alunos, estimou-se um modelo nulo em dois níveis para cada Unidade da Federação (UF) do país, isto é, foram ajustadas 27

10. Intuitivamente, a importância de se considerar uma análise em mais de um nível pode ser verificada no gráfico B.1, no apêndice B, que evidencia a variação da nota média de matemática por escola pública e por município no Brasil.

regressões de dois níveis do conjunto de equações realizadas logo acima. Assim, a partir dos resultados expostos no gráfico 1, podem-se observar os seguintes aspectos:

- com a divisão da amostra por estado, as diferenças de desempenho dos estudantes explicadas pelas escolas variam de localidade para localidade, não seguindo, no entanto, um padrão consistente com as diferenças de renda *per capita* desses estados;
- para o caso do Tocantins e do Piauí, a variabilidade das notas imputadas à unidade escolar foi acima de 20%, as mais altas do país, enquanto para o Distrito Federal e os estados de Rondônia e do Amapá, o percentual das diferenças das notas motivadas pelas distinções entre as escolas teve um poder de explicação de menos de 10% da variabilidade total das notas (gráfico 1); e
- como a literatura aponta que um VPC acima de 8% significa forte dependência hierárquica dos dados, salvo o caso do Amapá, um modelo de dois níveis seria justificável, inclusive, para uma análise por estado no Brasil.

**GRÁFICO 1**  
**Resultados do modelo de componentes de variância (modelo nulo) para os estados brasileiros nos níveis aluno-escola**  
 (Em %)



Elaboração dos autores.

Tendo em vista a grande heterogeneidade dos municípios brasileiros, em termos de características estruturais, sociais, econômicas, demográficas e territoriais, torna-se pertinente verificar a inclusão de um terceiro nível (município) para explicação dos diferentes resultados entre os alunos. Ou seja, quando se leva em conta o conjunto de fatores não observados entre as municipalidades do país, os próprios componentes aleatórios deste nível geram um melhor controle para o efeito das diversidades das notas médias entre municípios e da dependência hierárquica considerada.

O bloco de equações (9) representa um modelo nulo de três níveis: o primeiro inclui os estudantes ( $i = 1, \dots, I$ ), o segundo as escolas ( $j = 1, \dots, J$ ) e o terceiro os municípios ( $k = 1, \dots, K$ ). O resíduo agora é particionado dentro do componente do nível I ( $e_{0ijk}$ ), do componente do nível II ( $u_{0jk}$ ) e do componente do nível III ( $v_{0k}$ ).

$$y_{ijk} = \beta_{0ijk} x_{0ijk}$$

$$\beta_{0ijk} = \beta_0 + v_{0k} + u_{0jk} + e_{0ijk}$$

$$v_{0k} \sim N(0, \sigma_{v_0}^2)$$

$$u_{0jk} \sim N(0, \sigma_{u_0}^2)$$

$$e_{0ijk} \sim N(0, \sigma_{e_0}^2).$$

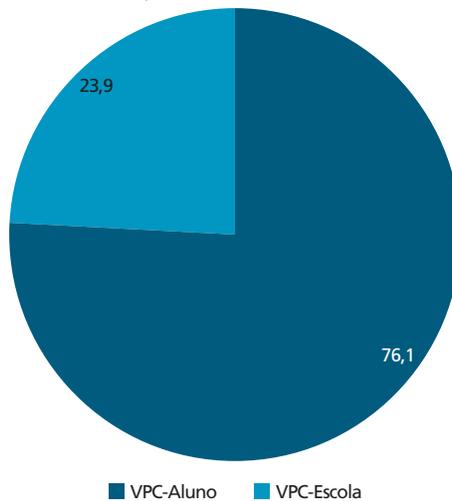
(9)

## GRÁFICO 2

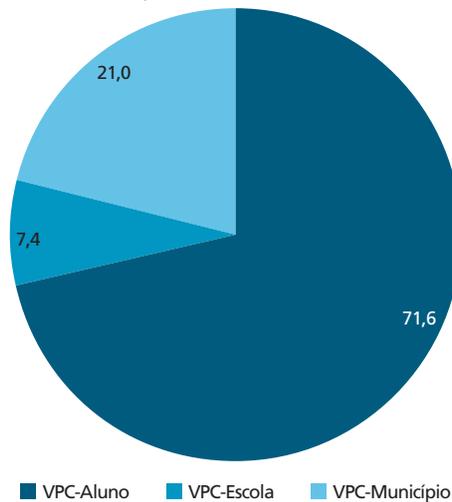
**Resultados do modelo nulo: Brasil**

(Em %)

2A – Resultados do modelo nulo para dois níveis



2B – Resultados do modelo nulo para três níveis



O gráfico 2B ilustra os resultados do VPC para um modelo de três níveis, que considera, além dos mais de 480 mil alunos e das mais de 20 mil escolas, os 4.311 municípios presentes na amostra selecionada neste estudo. No modelo de três níveis, os efeitos da nova estrutura hierárquica e da maior variabilidade das notas médias entre os municípios reduzem o percentual de explicação de variabilidade imputada às escolas, antes de 23,9% em um modelo de dois níveis (ver gráfico 2A), para 7,4% no modelo nulo com três níveis, ao passo que a heterogeneidade não observada entre os municípios explica 21,0% da variabilidade total das notas.

Ao comparar os resultados dos modelos nulos de dois níveis com os de três níveis, percebe-se que, no último, o papel da heterogeneidade municipal evidencia uma menor diversidade de notas entre as escolas do que entre o conjunto de municípios. Destarte, em termos gerais, a explicação conjunta da escola e do município em relação às diferentes notas auferidas pelos discentes da rede pública em matemática na Prova Brasil 2009 é superior à observada no modelo de dois níveis (aluno e escola).

Com o intuito de verificar o comportamento do VPC imputado à escola e ao município entre estados brasileiros, efetuaram-se 26 regressões em três níveis<sup>11</sup> (alunos, escolas e municípios), conforme (9). Deste modo, ao se analisarem as informações contidas no gráfico 3, observam-se as seguintes características:

- na maioria dos estados brasileiros, a variabilidade das notas entre os alunos tem mais relação com as diferenças entre as escolas do que pelas disparidades entre os municípios de cada estado.<sup>12</sup> Tal resultado informa que, em geral, existe menos heterogeneidade não observada entre as municipalidades em cada estado do que entre as próprias escolas. Assim, é válido reforçar que os resultados do gráfico 3 implicam que a variabilidade de notas entre municípios do mesmo estado é menor do que a variabilidade entre escolas. No entanto, quando se considera o Brasil, os municípios são comparados entre estados diferentes, o que amplifica a variabilidade;
- apenas para Goiás, Amapá e Rio Grande do Norte os efeitos da heterogeneidade dos municípios apresentaram um maior poder explicativo para a variação da proficiência em matemática dos estudantes do que os efeitos das escolas. Esse resultado, como se percebe, é uma

11. Para permitir variabilidade das notas por municípios em cada estado, os alunos das escolas do Distrito Federal foram agregados ao estado de Goiás.

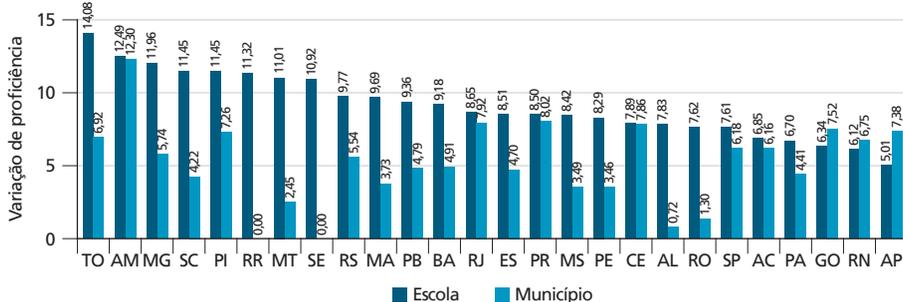
12. É válido lembrar que os resultados do gráfico 2 comparam a variabilidade entre todos os municípios do Brasil, o que aumenta a heterogeneidade deles. Em relação ao gráfico 3, comparam-se municípios de uma mesma UF.

particularidade quando comparado com os efeitos identificados para o restante dos estados brasileiros; e

- para os estados de Roraima e de Sergipe, a variação das notas explicadas pelos municípios foi nula, de modo que, para tais estados – para a amostra selecionada –, um modelo de três níveis não se justificaria. Neste sentido, parece haver pouca variabilidade entre o grupo considerado para o terceiro nível, de modo que bastaria apenas um modelo de dois níveis para uma análise desagregada para tais estados.<sup>13</sup>

### GRÁFICO 3

**Resultados do modelo de componentes de variância (modelo nulo) para os estados brasileiros nos níveis aluno-escola-município**  
(Em %)



Elaboração dos autores.

Em linhas gerais, os achados apresentados nesta subseção sugerem que a natureza hierárquica dos dados apresenta significativos efeitos contextuais sobre o desempenho individual de alunos, seja por parte de diferenças entre escolas, seja por distintas municipalidades.

#### 4.2 Descrição estatística das variáveis selecionadas na especificação econométrica

A partir da construção heurística do modelo, nem todas as variáveis pré-selecionadas e presentes no quadro 1 ajustaram-se estatisticamente ao modelo econométrico.

A tabela 1 apresenta as estatísticas descritivas para variáveis referentes aos alunos e às escolas de acordo com os dados da Prova Brasil que, conforme discutido adiante, melhor se adequaram à modelagem multinível. Em geral, pode-se caracterizar o aluno típico do 5º ano do ensino fundamental de escolas públicas/urbanas como uma criança do sexo masculino, de cor de pele não branca, que não apresenta reprovação e/ou abandono e que não trabalha. No entanto, ressalta-se que as

13. Isso ocorreu porque a variância estimada do termo aleatório do terceiro nível foi zero.

principais diferenças de proporção entre os alunos são registradas para o histórico de reprovação/abandono e o trabalho infantil.

A maioria dos alunos cursa a série adequada à sua idade, seus pais registram escolaridade média de 2,9, superior ao valor mediano (2,5) no índice de escolaridade. Há, contudo, elevado desvio-padrão nessas características médias, destacando fortes diferenças entre os atributos dos alunos.

**TABELA 1**  
**Estatísticas descritivas das variáveis selecionadas**

Variáveis	Média	Desvio-padrão	Mínimo	Máximo
<i>Alunos</i>				
<i>Gênero</i>				
Menina	0,4809	0,4996	0	1
Menino	0,5191	0,4996	0	1
<i>Raça</i>				
Não branco	0,6088	0,4880	0	1
Branco	0,3912	0,4880	0	1
<i>Reprovação e/ou abandono</i>				
Não	0,9640	0,1862	0	1
Sim	0,0360	0,1862	0	1
<i>Trabalha</i>				
Não	0,8708	0,3354	0	1
Sim	0,1292	0,3354	0	1
<i>Escolaridade dos pais</i>	2,9048	1,1359	1	5
<i>Defasagem idade-série</i>	0,6205	1,0600	-2	5
<i>Nota em matemática (em log)</i>	5,3439	0,2285	4,5362	5,8814
Observações (alunos)	481.725	-	-	-
<i>Escolas</i>				
Taxa de variação da nota em matemática (2007-2009)	0,1120	0,0943	-0,3479	0,8746
Proporção de reprovação e/ou de abandono (2009)	0,0506	0,0415	0	0,3333
Proporção de docentes efetivos	0,6828	0,3372	0	1
Observações (escolas)	20.545	-	-	-

Fonte: Prova Brasil de 2007 e 2009 (Inep, 2007; 2009b).

Entre as 20.545 escolas selecionadas na amostra, a média de crescimento do escore da escola em matemática entre 2007 e 2009 foi de 11,2%. A proporção média de alunos com histórico de reprovação e/ou abandono por escola em 2009 foi de 5,1%, enquanto a proporção média de professores efetivos foi 68,3%.

### 4.3 Evidências sobre os determinantes do desempenho escolar

A tabela 2 apresenta os resultados das regressões estimadas para a determinação do rendimento escolar dos alunos concluintes do 5º ano em escolas públicas no Brasil. Em todos os modelos estimados, variável dependente é a nota individual do aluno (em logaritmo) no exame de matemática realizado na Prova Brasil de 2009. Na coluna (1) encontram-se os resultados de referência, a partir da estimativa por mínimos quadrados ordinários (MQO), sem considerar qualquer heterogeneidade relacionada às variações de notas interescolas ou intermunicípios. Entretanto, os resultados obtidos usando a modelagem multinível (considerando a estrutura e a dependência hierárquica) são registrados nas demais colunas. Nas colunas de (2) a (4) foram feitas regressões lineares em dois níveis hierárquicos (alunos e escolas), com a entrada sucessiva de variáveis explicativas contextuais, enquanto a coluna (5) apresenta os resultados do modelo multinível em três níveis (alunos, escolas e municípios), controlando uma maior gama de diferenças não observadas a partir de efeitos aleatórios por município.

Comparando o ajuste de todas as regressões na tabela em destaque, ressaltasse que a regressão em três níveis (alunos, escolas e municípios), registrada na coluna (5), foi aquela que melhor se ajustou, considerando os menores valores das estatísticas de AIC e Deviance. É importante ressaltar que a especificação do modelo multinível segue a heurística *bottom-up*, em que geralmente se introduz, gradativamente, variáveis contextuais associadas a diferentes níveis superiores de hierarquia (segundo e terceiro níveis, por exemplo) até que se observe: *i*) redução das estatísticas supracitadas; e *ii*) correlação destas variáveis com o termo aleatório (componente do intercepto), o qual captura a heterogeneidade não observada entre os grupos. Neste sentido, entre as variáveis contextuais pré-selecionadas,<sup>14</sup> apenas as variáveis: *i*) variação percentual da nota média da escola em matemática na Prova Brasil de 2007 e de 2009 ( $x_1$ ); *ii*) proporção de alunos com histórico de reprovação e/ou abandono em 2009 ( $x_2$ ); e *iii*) proporção de professores efetivos em 2009 ( $x_3$ ) adequaram-se ao modelo.<sup>15</sup>

---

14. Ver quadro 1.

15. No âmbito do município, foi testada a inclusão da variável IFDM, de sua variação entre os anos de 2005 a 2009 e de 2007 a 2009. Todavia, além de não reduzir a Deviance, também foi registrada correlação com o componente aleatório do intercepto do modelo.

**TABELA 2**  
**Brasil: determinantes do rendimento escolar – regressões em três níveis hierárquicos**

	MQO		Multinível		
	(1)	(2)	(3)	(4)	(5)
Aluno					
Menino	0,0176*** (0,0006)	0,0180*** (0,0006)	0,0177*** (0,0006)	0,0177*** (0,0006)	0,0176*** (0,0006)
Branca	0,0214*** (0,0006)	0,0147*** (0,0006)	0,0134*** (0,0006)	0,0134*** (0,0006)	0,0130*** (0,0006)
Escolaridade dos pais	0,0180*** (0,0003)	0,0158*** (0,0003)	0,0151*** (0,0003)	0,0151*** (0,0003)	0,0159*** (0,0003)
Defasagem idade-série	-0,0332*** (0,0003)	-0,0329*** (0,0003)	-0,0317*** (0,0003)	-0,0317*** (0,0003)	-0,0314*** (0,0003)
Reprovação/abandono	-0,0594*** (0,0018)	-0,0690*** (0,0016)	-0,0636*** (0,0016)	-0,0636*** (0,0016)	-0,0634*** (0,0016)
Trabalha	-0,0768*** (0,0010)	-0,0698*** (0,0009)	-0,0691*** (0,0009)	-0,0691*** (0,0009)	-0,0688*** (0,0009)
Escola					
Varição da nota de matemática ( $x_1$ )	0,2949*** (0,0035)	0,3546*** (0,0081)	0,2958*** (0,0076)	0,2985*** (0,0076)	0,2569*** (0,0066)
Proporção reprovação/abandono ( $x_2$ )	-1,1637*** (0,0086)	-	-1,0999*** (0,0165)	-1,0869*** (0,0165)	-0,6227*** (0,0154)
Proporção professores efetivos ( $x_3$ )	0,0278*** (0,0010)	-	-	0,0256*** (0,0019)	0,0160*** (0,0017)
Intercepto ( $\beta_0$ )	5,3092*** (0,0014)	5,2645*** (0,0015)	5,3270*** (0,0017)	5,3087*** (0,0022)	5,2942*** (0,0024)
$var(x_1)$	-	0,0581*** (0,0090)	0,0795*** (0,0083)	0,0787*** (0,0083)	0,0410*** (0,0052)
$cov(x_1, x_2)$	-	-0,0056*** (0,0012)	0,0889*** (0,0135)	0,0907*** (0,0135)	0,0375*** (0,0089)
$var(x_2)$	-	-	0,2344*** (0,0378)	0,2249*** (0,0377)	0,1547*** (0,0257)
$cov(x_1, x_3)$	-	-	-0,0123*** (0,0015)	0,0012 (0,0016)	0,0001 (0,0010)
$cov(x_2, x_3)$	-	-	-0,0297*** (0,0035)	0,0037 (0,0033)	0,0026 (0,0022)
$var(x_3)$	-	-	-	0,0016** (0,0008)	0,0007 (0,0005)
$cov(x_1, \beta_1)$	-	-	-	-0,0133*** (0,0018)	-0,0056*** (0,0012)
$cov(x_2, \beta_1)$	-	-	-	-0,0309*** (0,0042)	-0,0169*** (0,0028)
$cov(x_3, \beta_1)$	-	-	-	-0,0011* (0,0006)	-0,0008** (0,0004)
$\sigma_u^2$ (escola)	-	0,0086*** (0,0002)	0,0088*** (0,0003)	0,0093*** (0,0006)	0,0043*** (0,0004)
$\sigma_e^2$ (aluno)	-	0,0376*** (0,0001)	0,0376*** (0,0001)	0,0376*** (0,0001)	0,0376*** (0,0001)
$\sigma_v^2$ (município)	-	-	-	-	0,0061*** (0,0002)
Observações	481.725	481.725	481.725	481.725	481.725
AIC	-138.441,12	-177.021,53	-181.208,09	-181.381,17	-188.074,05
Deviance	-	-177.045,53	-181.240,09	-181.423,17	-188.118,05

Elaboração dos autores.

Notas: \* Estatisticamente significante a 10%.

\*\* Estatisticamente significante a 5%.

\*\*\* Estatisticamente significante a 1%.

Obs.: Desvios-padrão robustos entre parênteses.

No tocante às covariadas associadas ao nível do aluno, os resultados mostram-se bem regulares em todas as regressões e os sinais dos coeficientes corroboram a expectativa prévia. Em linhas gerais, os meninos têm maior rendimento em matemática quando comparados às meninas (categoria omitida), assim como as crianças de cor branca apresentam, em média, melhor desempenho que aquelas não brancas (categoria-base). O trabalho infantil e o histórico individual de repetência ou de abandono dificultam o aprendizado dos estudantes. Note-se, ainda, que as crianças trabalhadoras registram menor resultado em matemática se comparadas aos alunos que não trabalham, por um lado, e, por outro, aquelas que já reprovaram ou abandonaram os estudos tiveram notas, em média, inferiores aos alunos sem relatos de reprovação e/ou abandono. O aluno com atraso escolar, capturado a partir de um maior valor para a variável *distorção idade-série*, está condicionado a um desempenho médio inferior, ou seja, para cada ano de defasagem na série, a nota média no exame de matemática diminui em cerca de 3 pontos de *log*. Já a escolaridade dos pais relaciona-se positivamente com o desempenho do estudante, aspecto possivelmente atrelado à melhor condição socioeconômica no domicílio (juntamente à raça) e/ou à maior atenção dos pais aos estudos dos filhos, que pode favorecer a transmissão e a persistência intergeracional de capital humano. Desta forma, para cada ano adicional na escolaridade dos pais, o rendimento escolar da criança em matemática aumenta em cerca de 1,5 ponto de *log*.

Os resultados também mostram que o papel do contexto escolar não pode ser desprezado quando se analisa a determinação do aprendizado do aluno. As variáveis da escola que mais contribuíram para o grau de ajustamento das regressões e se revelaram correlacionadas à heterogeneidade não observada entre escolas foram: variação percentual da nota média da escola em matemática na Prova Brasil de 2007 e de 2009 ( $x_1$ ); proporção de alunos com histórico de reprovação e/ou abandono em 2009 ( $x_2$ ); e proporção de professores efetivos em 2009 ( $x_3$ ). É bastante razoável admitir que esse conjunto de covariadas capture diferenças entre a qualidade das escolas e/ou diferenças de eficácia da gestão escolar no alcance de melhores resultados. Neste sentido, os resultados da tabela 2 sugerem que os alunos com melhor aprendizado são aqueles que estudam em escolas que registraram maior crescimento da nota média entre 2007 a 2009, assim como menores taxas de reprovação/abandono e elevadas taxas de professores efetivos. Esses achados mantêm-se regulares quando se controla a heterogeneidade não observada entre os municípios, conforme os resultados registrados na coluna (5). Tais evidências sugerem que escolas públicas com melhor qualidade ou onde há maior esforço de gestão no alcance de resultados escolares positivos explicam boa parte das diferenças de rendimento entre os alunos da primeira fase do ensino fundamental no Brasil.

TABELA 3

**Brasil: determinantes do rendimento escolar – efeitos dos fatores contextuais da escola: regressões em três níveis hierárquicos por estados**

Estado	Variação da nota de matemática (2007-2009)	Proporção reprovação/abandono	Proporção professores efetivos	Municípios (nível 3)	Escolas (nível 2)	Alunos (nível 1)
Rondônia	0,2587*** (0,0692)	-0,4260*** (0,1493)	0,0041 (0,0138)	42	163	3.326
Acre	0,3861*** (0,1108)	-0,3637 (0,2477)	0,0038 (0,0217)	19	79	1.566
Amazonas	0,3269*** (0,0465)	-0,2965*** (0,1139)	-0,0128 (0,0138)	58	444	10.496
Roraima	0,1475 (0,1316)	-0,8843*** (0,2772)	0,0089 (0,0275)	5	58	1.134
Pará	0,2729*** (0,0257)	-0,2734*** (0,0480)	0,0026 (0,0060)	139	932	19.921
Amapá	0,2426*** (0,0771)	-0,1870 (0,1823)	-0,0351* (0,0202)	12	105	2.239
Tocantins	0,3805*** (0,0771)	-0,5840*** (0,1921)	-0,0156 (0,0180)	82	168	2.786
Minas Gerais	0,1769*** (0,0166)	-0,8049*** (0,0494)	0,0165*** (0,0048)	682	2.140	53.673
Espírito Santo	0,1781*** (0,0395)	-0,5997*** (0,1063)	0,0063 (0,0094)	73	433	8.106
Rio de Janeiro	0,1997*** (0,0244)	-0,3661*** (0,0586)	-0,0028 (0,0074)	81	1375	26.710
São Paulo	0,2224*** (0,0137)	-0,8033*** (0,0372)	0,0196*** (0,0033)	573	4.434	150.029
Maranhão	0,2520*** (0,0323)	-0,4061*** (0,0710)	-0,0062 (0,0090)	180	594	10.422
Piauí	0,2862*** (0,0555)	-0,5066*** (0,0943)	0,0034 (0,0139)	106	261	4.678
Ceará	0,2392*** (0,0342)	-0,4251*** (0,0692)	-0,0112 (0,0094)	153	655	12.587
Rio Grande do Norte	0,2299*** (0,0452)	-0,2247*** (0,0941)	0,0298*** (0,0143)	105	263	4.590
Paraíba	0,2707*** (0,0446)	-0,3456*** (0,0881)	0,0116 (0,0112)	136	363	6.004
Pernambuco	0,3148*** (0,0305)	-0,2241*** (0,0664)	0,0108 (0,0081)	172	733	13.171
Alagoas	0,0793 (0,1025)	-0,4465*** (0,1587)	0,0370 (0,0287)	33	61	1.217
Sergipe	0,2282*** (0,0608)	-0,3364*** (0,1097)	0,0132 (0,0173)	57	203	3.569
Bahia	0,3186*** (0,0276)	-0,3375*** (0,0460)	0,0088 (0,0067)	351	1.177	21.601
Paraná	0,3659*** (0,0235)	-0,5650*** (0,0626)	0,0170*** (0,0059)	375	1.682	41.515
Santa Catarina	0,2752*** (0,0311)	-0,6617*** (0,0925)	0,0312*** (0,0080)	230	1.004	21.156
Rio Grande do Sul	0,2639*** (0,0240)	-0,5532*** (0,0636)	0,0098 (0,0062)	293	1.521	25.476
Mato Grosso do Sul	0,1590*** (0,0454)	-0,3528*** (0,1182)	0,0392*** (0,0111)	72	376	7.284
Mato Grosso	0,2931*** (0,0475)	-0,4403*** (0,1125)	-0,0061 (0,0113)	96	374	7.369
Goiás/DF	0,2813*** (0,0281)	-0,3630*** (0,0704)	0,0337*** (0,0065)	186	947	21.100

Elaboração dos autores.

Notas: \* Estatisticamente significante a 10%.

\*\* Estatisticamente significante a 5%.

\*\*\* Estatisticamente significante a 1%.

Obs.: Desvios-padrão robustos entre parênteses.

Considerando que avanços no desempenho médio da escola entre duas avaliações pode ser o resultado do esforço empreendido e da melhoria nos insumos escolares, pode-se inferir, indiretamente, que os resultados encontrados atestam a efetividade do gasto público destinado às unidades de ensino, pois, em última instância, a grande meta dos recursos públicos em educação é o fornecimento de educação de qualidade. Além disso, aquelas unidades escolares que apresentaram uma maior proporção de docentes em situações estáveis (concursados), mais uma nítida variável de preocupação da gestão educacional em auferir melhores resultados, mostrou-se ser importante para a obtenção de melhores escores para os estudantes.

Ao se comparar os resultados estimados por MQO (1) e com aqueles produzidos pelo modelo multinível (5), nota-se que a contribuição marginal de todas as variáveis contextuais da escola, em termos absolutos, é superestimada pela abordagem tradicional (ingênua). Em relação às variáveis dos alunos, ora as estimativas por MQO subestimam ora superestimam os valores de tais coeficientes.

Na tentativa de se explorar melhor possíveis diferenças regionais em termos das variáveis contextuais às escolas, a saber, diferença por UFs devido a regimes locais específicos, foram estimados modelos de três níveis (alunos, escolas e municípios) em amostras separadas por estados,<sup>16</sup> conforme especificação final<sup>17</sup> na coluna (5) da tabela 2. Os resultados são elencados na tabela 3, em que são registrados apenas os coeficientes estimados para as variáveis contextuais das escolas.<sup>18</sup>

Em linhas gerais, os resultados são consistentes com aqueles observados nas regressões feitas para a amostra total no Brasil, dada a regularidade dos sinais e a consonância com a expectativa de sinais. Contudo, algumas diferenças regionais podem ser destacadas no tocante aos efeitos marginais. Nos estados do Acre, do Tocantins, do Paraná, do Amazonas, da Bahia e de Pernambuco, as correlações positivas entre a melhoria da nota média da escola e o desempenho individual do aluno são relativamente superiores àquelas registradas para as demais UFs. Já em relação à taxa de reprovação e/ou abandono, os estados que apresentaram correlações relativamente superiores foram, respectivamente, Roraima, Minas Gerais, São Paulo e Santa Catarina, onde escolas públicas com maior nível de reprovação/abandono registram, em média, estudantes com baixas notas individuais. Por fim, destacam-se os estados de Santa Catarina, de Mato Grosso do Sul e de Goiás/DF, onde as escolas com maior proporção de professores efetivos registraram uma associação direta e relativamente superior aos demais estados no que concerne às notas individuais dos alunos concluintes da primeira fase do ensino fundamental.

16. Para permitir variabilidade entre os municípios, agregou-se os alunos e escolas do Distrito Federal ao estado de Goiás.

17. Essa estratégia se mostrou menos custosa e mais parcimoniosa que a introdução de um quarto nível hierárquico no modelo.

18. Os resultados gerais podem ser prontamente disponibilizados pelos autores.

## 5 CONSIDERAÇÕES FINAIS

Este trabalho procurou contribuir para o melhor entendimento dos determinantes do desempenho escolar no Brasil, ao estimar uma função de produção escolar condicionada a uma estrutura hierárquica de dados com covariadas socioeconômicas do aluno, variáveis contextuais da escola e diferenças não observadas entre as municipalidades. Neste sentido, buscou-se avançar no controle para variáveis usualmente omitidas em estudos da literatura especializada, particularmente variáveis associadas ao esforço da gestão escolar no tocante à melhoria da qualidade de ensino.

A análise de decomposição de variância em modelos nulos revelou que as diferenças entre os alunos preponderam sobre as diferenças entre escolas e municípios na explicação da variabilidade das notas na prova de matemática. Ao se considerar hierarquias em níveis aluno-escola-município, os achados mostraram que notas individuais variam mais entre municípios do que entre escolas quando se incorporam diferenças interestaduais. No ambiente intraestadual, as diferenças contextuais entre as escolas parecem ser mais relevantes na variabilidade das notas.

Todavia, o principal achado deste estudo revela a importância de alguns fatores contextuais da escola no desempenho do aluno. Ou seja, os achados mostraram que a nota do aluno na prova de matemática relaciona-se diretamente com a elevação do desempenho médio da escola nos anos anteriores e com a maior presença de professores efetivos no quadro docente, por um lado, e, por outro, associa-se inversamente com o histórico de reprovação e/ou abandono na escola. Tais resultados persistem mesmo após se controlar as regressões para diferenças não observadas em municipalidades. Uma vez que tais variáveis contextuais, provavelmente, captam o esforço da gestão escolar no objetivo da orientação de uma oferta de serviços educacionais com melhor qualidade, este estudo mostra, portanto, que diferenças em gestão escolar podem alterar o desempenho escolar dos estudantes.

Outros achados derivados da estimação da função de produção corroboram a literatura nacional sobre o tema, tais como o melhor desempenho dos meninos e das crianças brancas no exame de matemática, comparado às meninas e às crianças não brancas, respectivamente. A escolaridade dos pais afeta positivamente o desempenho dos alunos, enquanto a defasagem idade-série, a reprovação/abandono e o fato de a criança trabalhar associam-se negativamente com a métrica de aprendizado.

No tocante à formulação de políticas públicas, os resultados deste estudo sugerem que a substituição de professores temporários por professores efetivos pode gerar um efeito positivo sobre o aprendizado do aluno. Políticas de geração de incentivos aos diretores/professores e de combate à reprovação e ao abandono escolar podem atuar na mesma direção. Portanto, espera-se que os achados reportados sirvam de referência para os formuladores de políticas públicas no que tange à oferta dos serviços educacionais com maior qualidade.

## REFERÊNCIAS

- ADDONIZIO, M. X-efficiency and effective schools: a new look at old theories. **Journal of Education Finance**, v. 35, n. 1, p. 1-25, 2009.
- ALMEIDA, A. T. C. Determinantes dos piores e melhores resultados educacionais dos alunos da rede pública de ensino fundamental no Brasil. **Planejamento e Políticas Públicas**, n. 42, p. 147-187, 2014.
- ANDRADE, J.; LAROS, J. Fatores associados ao desempenho escolar: estudo multinível com dados do Saeb/2001. **Psicologia: teoria e pesquisa**, v. 23, n. 1 p. 33-42, 2007.
- BAFUMI, J.; GELMAN, A. E. Fitting multilevel models when predictors and group effects correlate. *In*: ANNUAL MEETING OF THE MIDWEST POLITICAL SCIENCE ASSOCIATION, 2006, Chicago. **Annals...** Chicago: MPSA, 2006.
- BARRO, R. J. Human capital and growth. **American Economic Review**, v. 91, n. 2, p. 12-17, 2001.
- BARROS, A. R. **Desigualdades regionais no Brasil: natureza, causas, origem e soluções**. Rio de Janeiro: Elsevier, 2011.
- BRASIL. Secretaria do Tesouro Nacional. **Finanças do Brasil: dados contábeis dos municípios**. Brasília: STN, 2010.
- BRETON, T. The quality vs. the quantity of schooling: what drives economic growth? **Economic Education Review**, v. 30, p. 765-773, 2011.
- CENEVIVA, R. **O nível de governo importa para a qualidade da política pública? O caso da educação fundamental no Brasil**. 2011. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2011.
- CLOTFELTER, C. T.; LADD, H. F.; VIGDOR, J. L. **How and why do teaching credentials matter for student achievement?** Cambridge: NBER, 2007. (NBER Working Paper, n. 12828).
- COLEMAN, J. *et al.* **Equality of educational opportunity**. Washington: US Government Printing Office, 1966.
- COORAY, A. V. The role of education in economic growth. *In*: AUSTRALIAN CONFERENCE OF ECONOMISTS, 2009, Adelaide. **Annals...** Adelaide: South Australian Branch of the Economic Society of Australia, 2009.
- COSTA, L. O.; ARRAES, R. A.; GUIMARÃES, D. B. Estabilidade dos professores e qualidade do ensino de escolas públicas. **Economia Aplicada**, v. 19, n. 2, p. 261-298, 2015.

DIAZ, M. D. M. Qualidade do gasto público em educação no Brasil. *In: FIPE – FUNDAÇÃO INSTITUTO DE PESQUISAS ECONÔMICAS* (Org.). **Qualidade do gasto público no Brasil: sugestões para melhorar os resultados das políticas públicas, sem aumento de impostos.** São Paulo: Fipe, 2007.

FLETCHER, P. R. À procura do ensino eficaz. Brasília: MEC, 1998.

FRANCO, A. M. P. **Os determinantes da qualidade da educação no Brasil.** 2008. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2008.

FRANCO, C.; MANDARINO, M.; ORTIGÃO, M. I. O projeto pedagógico e os resultados escolares. **Pesquisa e Planejamento Econômico**, v. 32, n. 3, p. 477-494, 2002.

GELMAN, A.; HILL, J. **Data analysis using regression and multilevel/hierarchical models.** New York: Cambridge University Press, 2007.

HANUSHEK, E. A. Interpreting recent research on schooling in developing countries. **The World Bank Research Observer**, v. 10, n. 2, p. 227-246, 1995.

\_\_\_\_\_. Some U.S. evidence on how the distribution of educational outcomes can be changed. *In: WOESSMANN, L.; PETERSON, P. E. Schools and the equal opportunity problem.* Cambridge: MIT Press, 2007.

\_\_\_\_\_. Education production functions: developed country evidence. *In: PETERSON, P.; BAKER, E.; MCGAW, B. (Eds.). International Encyclopedia of Education.* Oxford: Elsevier, 2010. v. 2, p. 407-411.

HANUSHEK, E. A.; KIMKO, D. D. Schooling, labor-force quality and the growth of nations. **American Economic Review**, v. 90, n. 5, p. 1184-1208, 2000.

HANUSHEK, E. A.; WOESSMANN, L. The role of cognitive skills in economic development. **Journal of Economic Literature**, v. 46, n. 3, p. 607-668, 2008.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo Demográfico 2000.** Rio de Janeiro: IBGE, 2000.

\_\_\_\_\_. **Censo Demográfico 2010.** Rio de Janeiro: IBGE, 2010.

INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados Prova Brasil 2007.** Brasília: Inep, 2007.

\_\_\_\_\_. **Índice de Desenvolvimento da Educação Básica 2009.** Brasília: Inep, 2009a.

\_\_\_\_\_. **Microdados Prova Brasil 2009.** Brasília: Inep, 2009b.

LEME, M. C.; PAREDES, R.; SOUZA, A. P. A municipalização do ensino fundamental e seu impacto sobre a proficiência no Brasil. *In: VELOSO, F. et al. (Org.). Educação básica no Brasil: construindo o país do futuro.* Rio de Janeiro: Campus, 2009.

MENEZES-FILHO, N. **Os determinantes do desempenho escolar do Brasil.** São Paulo: TPE, 2007. Disponível em: <<http://goo.gl/wsXqBF>>.

NASCIMENTO, P. A. M. M. Desempenho escolar e gastos municipais por aluno em educação: relação observada em municípios baianos para o ano 2000. **Ensaio: avaliação e políticas públicas em educação**, v. 15, n. 56, p. 393-412, 2007.

NATIS, L. Modelos lineares hierárquicos. **Estudos em Avaliação Educacional**, n. 23, p. 3-29, 2001.

OATES, W. E. **Federalismo fiscal.** Madri: Instituto de Estudios de Administración Local, 1977.

\_\_\_\_\_. Toward a second-generation theory of fiscal federalism. **International Tax and Public Finance**, v. 12, n. 4, p. 349-373, 2005.

OECD – ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Pisa 2009 results: executive summary.** Washington: OECD, 2010.

PINHEIRO, J. C.; BATES, D. M. **Mixed-effects models in S and S-Plus.** New York: Springer, 2000.

RASBASH, J. *et al.* **A user's guide to MLwiN.** Bristol: University of Bristol, 2009.

RIANI, J. L. R.; RIOS-NETO, E. L. G. Background familiar versus perfil escolar do município: qual possui maior impacto no resultado educacional dos alunos brasileiros? **Revista Brasileira de Estudos Populacionais**, v. 25, n. 2, p. 251-269, 2008.

RIVKIN, S. G.; HANUSHEK, E. A.; KAIN, J. F. Teachers, schools, and academic achievement. **Econometrica**, v. 73, n. 2, p. 417-458, 2005.

SOARES, J. F.; ALVES, M. T. G. Efeitos de escolas e municípios na qualidade do ensino fundamental. **Cadernos de Pesquisa**, v. 43, n. 149, p. 492-517, 2013.

SOARES, J. F.; CANDIAN, J. F. O efeito da escola básica brasileira: as evidências do Pisa e do Saeb. **Revista Contemporânea de Educação**, v. 2, n. 4, p. 1-12, 2007.

SOARES, T. M. Influência do professor e do ambiente em sala de aula sobre a proficiência alcançada pelos alunos avaliados no Simave-2002. **Estudos em Avaliação Educacional**, n. 28, p. 103-123, 2003.

SOLMON, L. C. Education quality and economic growth. **Economics of Education Review**, v. 4, n. 4, p. 273-290, 1985.

## APÊNDICE A

TABELA A.1  
**Estatística descritiva das variáveis pré-selecionadas ao modelo multinível para os níveis aluno e escola**

Variável	Média	Desvio-padrão
<i>LNOTAM</i>	5,344	0,228
<i>SEXO</i>	0,519	0,500
<i>REPROVACAO</i>	0,036	0,186
<i>DFIDADE</i>	0,621	1,060
<i>TRABALHA</i>	0,129	0,335
<i>RACA</i>	0,391	0,488
<i>PAIS_ESTUDO</i>	2,905	1,136
<i>T_SITTRAB</i>	0,687	0,317
<i>ESC_PEDAPORT</i>	0,861	0,188
<i>ESC_PROFESSORES</i>	3,639	1,978
<i>ESC_SALESCOLA</i>	0,463	0,177
<i>ENOTAM</i>	215,865	23,004
<i>DELTA_ENOTAM</i>	0,110	0,089
<i>SNOTAM</i>	43,127	5,607
<i>TDI</i>	0,478	0,179
<i>T_REPROVACAO</i>	0,046	0,038
<i>T_TRABALHA</i>	0,120	0,062
<i>ESC_PAIS_ESTUDO</i>	2,906	0,440

Observações: 481.725

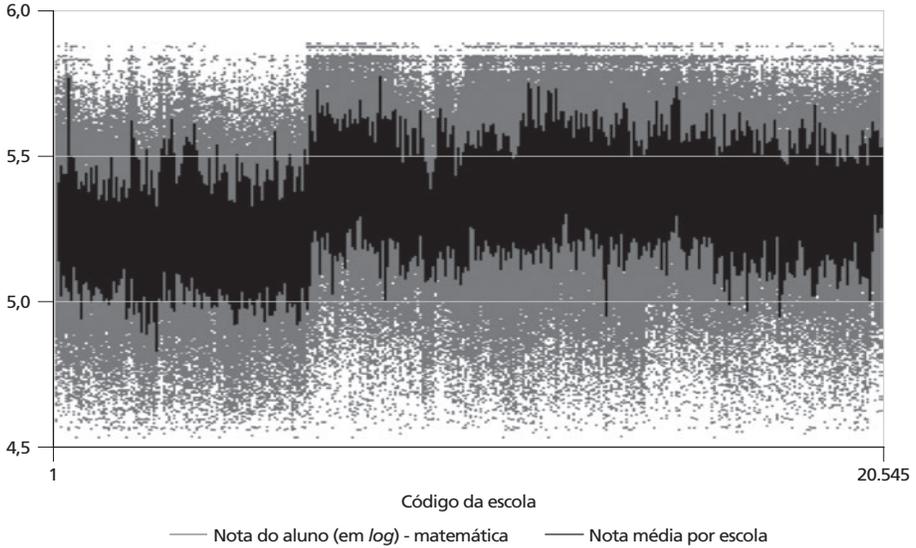
Elaboração dos autores.

## APÊNDICE B

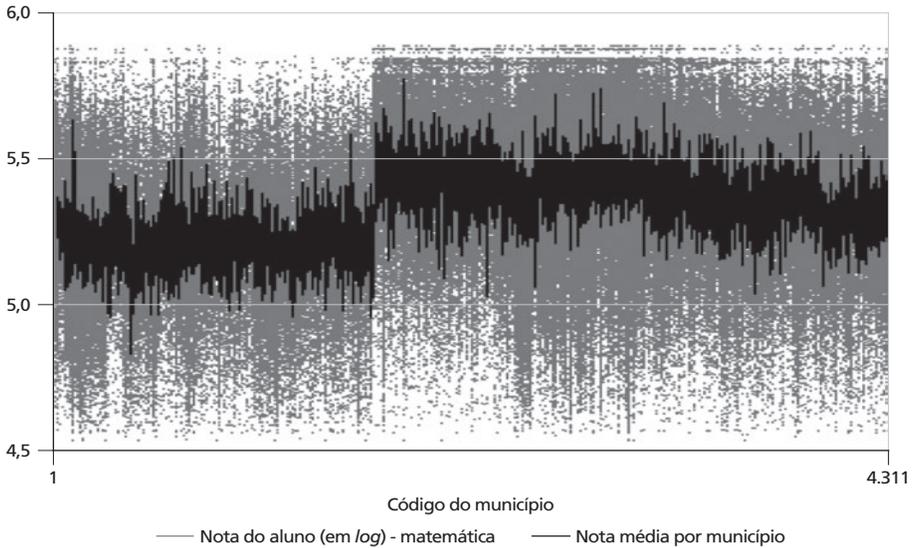
### GRÁFICO B.1

**Brasil: variação da nota média de matemática por escola pública e por município – alunos concluintes da primeira fase do ensino fundamental**

B.1A – Nota média por escola



B.1B – Nota média por município



Elaboração dos autores.