

# Qualidade do Ensino em Matemática: Determinantes do Desempenho de Alunos em Escolas Públicas Estaduais Mineiras

Ana Flávia Machado, Sueli Moro,  
Ludiemy Martins, Juan Rios

*Universidade Federal de Minas Gerais (CEDEPLAR-UFMG), Minas Gerais, Brasil*

---

## Resumo

Este artigo analisa os determinantes do desempenho de alunos em Matemática nas séries finais do primeiro e segundo ciclo do Ensino Fundamental e do Ensino Médio em Minas Gerais no ano de 2003. Modelos hierárquicos em três níveis são utilizados. No primeiro nível, nós incluímos muitos atributos de alunos e de seu background familiar; no segundo, algumas características das escolas, principalmente o culto-aluno; no terceiro, variáveis referentes ao município onde a escola se localiza. Quatro fontes de dados são utilizadas: Sistema de Avaliação de Ensino (SIMAVE), Sistema de Informação Custo Aluno (SICA), Censo Escolar e o de Indicadores de Responsabilidade Social de Minas Gerais (IMRS), todos para 2003. Os resultados mostram que, para todas as séries, a escolaridade da mãe e a presença de, no mínimo vinte livros em casa, contribuem positivamente para o desempenho do aluno em Matemática. Para o Ensino fundamental, é importante, também, realizar a tarefa de matemática em casa. Para a 3ª série do Ensino Médio, não alocar tempo para tarefas domésticas contribui positivamente para o resultado nos testes de Matemática.

*Palavras-chave:* Educação, Qualidade, Modelos Hierárquicos

*Classificação JEL:* I20, I21

---

## Abstract

This article analyses the determinants of the performance in mathematics of pupils in the Elementary School and High School in Minas Gerais, Brazil, in 2003. Hierarchical models in three levels were used. In the first level we included several attributes of the pupils and their familiar background; in the second level some attributes of the school, mainly the cost-pupil and in the third level, some characteristics of the municipalities where the Schools are located. Four databases were used in this study, the System of Evaluation of Public Education for Minas Gerais State (SIMAVE), the Information System for Costs per Pupil (SICA), the School Census and the Index of Social Responsibility for the State of Minas Gerais (IMRS), all for 2003. We point out, for all grades the positive effects of pupils' academic performance of mother's education and the presence of at least, some books at home. For Elementary School, besides these

effects, the positive impact of doing the math homework at home was important. For the third grade of High School, not to allocate time for household tasks was also an important explanatory variable.

---

## 1. Introdução

Os indicadores referentes à educação, no Brasil, tais como taxa de analfabetismo, taxa de atendimento escolar, taxa de escolarização líquida e, em decorrência, o nível de escolaridade média da população, melhoraram bastante nas últimas três décadas, embora ainda estejam aquém do observado em países de desenvolvimento similar.

Neste contexto, emergem questões referentes à eficiência da provisão de serviços em educação. Segundo Afonso e Aubyn (2005), um sistema educacional pode ser considerado eficiente, como qualquer outra atividade, se o acréscimo no produto se faz às expensas de um gasto com insumos constante ou se obtém um mesmo nível de produto com um gasto em insumo menor. À primeira vista, esta definição clássica da teoria econômica é de difícil aplicação à atividade de ensino, porque educação formal não é um bem qualquer. O que vem a ser o produto? Certamente, baixa taxa de distorção idade-série, elevada taxa de aprovação e baixa de reprovação e evasão, proficiência em disciplinas-chave dos currículos são alguns dos produtos. Por outro lado, quais são os insumos? Professores qualificados, infra-estrutura escolar adequada, como presença de salas de aulas, carteiras, quadro-negro e biblioteca, laboratórios, equipamentos de informática, etc. Esta análise não é trivial, mas é de suma importância para um país como o nosso que se encontra em um processo de universalização da educação e, assim, como a maioria dos países, tem o Estado como seu maior provedor.

Este artigo busca, portanto, analisar os determinantes do desempenho de alunos do ensino básico da rede pública estadual mineira na disciplina Matemática em 2003. Por que Minas Gerais? Em primeiro lugar, porque o estado de Minas Gerais reflete o contexto nacional, dado que o estado, tal como Brasil, é um caleidoscópio de culturas, costumes, atividades econômicas e estágios de desenvolvimento. Ao mesmo tempo, em que se vê pobreza e miséria no Jequitinhonha e Vale do Mucuri, têm-se, por exemplo, a agricultura mecanizada na produção de grãos do Triângulo Mineiro, a indústria pesada do Vale do Aço, as empresas de tecnologia do sul de Minas e os serviços modernos da Região Metropolitana. Logo, não por acaso, construir indicadores/medidas de educação a partir da realidade mineira é ter um instrumento válido para aplicá-lo à imensa diversidade do Brasil. Em segundo lugar, tem-se uma base de dados de custo-aluno de escolas públicas estaduais, algo ímpar na experiência nacional. A Secretária de Educação desenvolveu um sistema

---

\* Recebido em maio de 2007, aprovado em novembro de 2007. Agradecemos aos comentários e sugestões dos pareceristas e a Sergei Soares pelas informações sobre o método de imputação de *missing*. Entretanto, erros e omissões são de responsabilidade dos autores.  
E-mail address: afmachad@cedepiar.ufmg.br.

de informações sobre o custo-aluno em caráter censitário que permite avaliar a eficiência da provisão de serviços em educação no âmbito das escolas do estado.

Assim, por intermédio de modelos hierárquicos, avaliamos, em um primeiro nível, as condições do aluno e de seu *background* familiar, em um segundo nível, as características socioeconômicas da escola, incluindo o custo-aluno e, em um terceiro nível, o município onde reside. Estimativas semelhantes, recorrendo a modelos hierárquicos, já foram feitas por Fletcher (1998), César e Soares (2001), Barbosa e Fernandes (2001), Ferrão et alii (2001), Albernaz et alii (2002) e Riani (2005), entre outros. Este artigo, contudo, traz a novidade de combinar bases de caráter censitário como o Sistema Mineiro de Avaliação da Educação Pública (SIMAVE), o Sistema Informacional Custo Aluno (SICA)TP<sup>1</sup> PT, além do Censo Escolar, organizado pelo INEP, mas, sobretudo, a de incluir informações referentes ao custo da educação como um dos determinantes da proficiência, pois uma das variáveis-chave em análise de eficiência é o custo.

O trabalho está dividido em cinco seções, além desta Introdução. Na segunda seção, revisamos, de forma sucinta, a literatura sobre o tema, especialmente a nacional. O modelo hierárquico é apresentado na terceira seção. Em seguida, na quarta, descrevemos as fontes de dados e o tratamento das variáveis, assim como a análise descritiva. A quinta seção traz os resultados da estimação do modelo hierárquico para as séries 4<sup>a</sup> e 8<sup>a</sup> do Ensino Fundamental e 3<sup>a</sup> do Ensino Médio. Por fim, são feitas algumas considerações.

## 2. Desempenho Escolar e seus Determinantes: Breve Revisão da Literatura

De uma maneira geral, considera-se uma “função de produção educacional”, explicando o desempenho dos alunos a partir de características dos estudantes e seu *background* familiar, ou seja, os aspectos pessoais e socioeconômicos do aluno, além de insumos escolares:

$$Y = F(cB_a, cB_f, cB_e, cB_pB, cB_m, \epsilon)$$

$Y$  – desempenho dos alunos medido pela proficiência;

$cB_a$  – vetor de características dos alunos tais como cor, sexo, idade;

$cB_f$  – vetor de características de sua família tais como escolaridade dos pais, renda e outras medidas de seu nível socioeconômico como disponibilidade de livros, acesso a equipamentos como computadores, trabalho da criança dentro ou fora do lar;

$cB_e$  – vetor de características da escola tais como escolaridade dos professores, infra-estrutura da escola, custo-aluno da escola, entre outros insumos;

$cB_m$  – vetor de características dos municípios onde a escola se localiza tais como PIB per capita municipal e indicadores de oferta de educação;

$\epsilon$  – termo de erro aleatório.

<sup>1</sup> As bases SIMAVE e SICA são explicadas na quarta seção referente à fonte de dados e tratamento das variáveis.

No âmbito de trabalhos que utilizam outras metodologias que não os modelos hierárquicos, Felício e Fernandes (2005) argumentam que, em regra, o efeito da escola é reduzido quando comparado à parcela explicada pelo *background* familiar; entretanto, essa relevância escolar pode estar sendo subestimada pela escolha dos insumos escolares.

Silva e Hasenbalg (2001), analisando três dimensões – recursos econômicos disponíveis para os gastos educacionais; recursos educacionais ou capital cultural da família e a estrutura dos arranjos familiares – mostram, por intermédio da PNAD de 1999, que o impacto das variáveis de *background* familiar é maior até a metade do ensino fundamental.

O trabalho de Barros et alii (2001), ao abranger as regiões Nordeste e Sudeste com base na PNAD de 1996 e na Pesquisa de Padrão de Vida (PPV) de 1996/97, evidencia que o fator mais importante para explicar o desempenho educacional é a escolaridade dos pais (principalmente a da mãe). Além disso, a qualidade de infra-estrutura da escola (mais importante na segunda etapa do ensino fundamental, diminuindo no ensino médio) tem efeito tão relevante quanto, ou maior que, a escolaridade dos professores.

Por outro lado, Hanushek et alii (1996), analisando a área rural do Nordeste, evidenciam que abastecimento de água e energia elétrica, instalações sanitárias, mobiliário para estudantes e professores, livros textos, guias para professores, recursos audiovisuais, *notebooks* e materiais de escritório estão relacionados positivamente com o desempenho estudantil. Por outro lado, variáveis ligadas à qualidade dos professores, como testes de desempenho de professores e participação em programas específicos de treinamento, não demonstraram relevância na explicação do desempenho escolar.

Considerando os estudos que recorrem a modelos hierárquicos, Albernaz et alii (2002) justificam a aplicação desta classe de modelos, porque acreditam que, na estimação por Mínimos Quadrados Ordinários de modelos de um só nível, os efeitos das variáveis de escola ficam subestimados *vis-à-vis* os efeitos do nível socioeconômico das famílias. Utilizando os dados do SAEB (Sistema de Avaliação do Ensino Básico) de 1999 para 8ª série, os autores estimam, em primeiro lugar, o modelo incondicional que representa o desempenho do aluno como a soma do desempenho médio da escola em matemática, variáveis *dummies* do desempenho do aluno em outras matérias, e de um resíduo que depende do aluno e da escola. Depois, incluem num segundo nível, a escola, com informações que descrevem o Nível Socioeconômico médio de cada Estabelecimento (NSE) explicando o desempenho médio de cada um destes e o nível socioeconômico dos alunos da respectiva escola (o coeficiente desta variável muda de acordo com o nível socioeconômico médio da escola). Os resultados evidenciam que o valor médio da variável NSE em cada escola está relacionado com o desempenho médio do aluno. No primeiro nível, o do aluno, seu próprio NSE contribui de forma significativa para o seu desempenho, mas em menor grau. Ademais, elevado nível socioeconômico médio da escola tende a suavizar o impacto negativo no desempenho causado por um baixo nível socioeconômico do aluno, propiciando maior equidade.

Albernaz, Ferreira e Franco ainda incluem, no segundo nível, variáveis como infra-estrutura da escola, recursos financeiros, e práticas escolares. A introdução desses novos fatores cancelou parte dos efeitos antes atribuídos à rede de ensino e ao nível socioeconômico médio de seus alunos. A proporção de repetentes teve um impacto grande e significativo no desempenho médio das escolas, mesmo controlando pelo nível socioeconômico. Os três grupos de variáveis escolares foram relevantes, sendo que as relacionadas aos recursos financeiros apresentaram os maiores parâmetros. Em termos da equidade, o modelo de Albernaz, Ferreira e Franco revelou que um maior NSE médio por escola faz com que o NSE de cada aluno afete menos sua proficiência (como no primeiro modelo), e um maior nível de escolaridade do professor tem um efeito oposto, a equidade aumenta.

Há, no entanto, aplicações de modelos hierárquicos que corroboram o papel do *background* familiar. Analisando as probabilidades de progressão da 1ª e 5ª séries, Rios-Neto et alii (2002) TP<sup>2</sup> PT utilizam, no primeiro nível, o indivíduo e, no segundo nível, a situação do domicílio e o tipo de área (metropolitana e não-metropolitana). No primeiro nível, os atributos do indivíduo são descritos pela educação da mãe e variáveis indicadoras de ocupação e posição na ocupação do pai. No segundo nível, as variáveis de controle foram anos médios de estudo, salário médio dos professores do ensino fundamental e razão professores do ensino fundamental/população em idade escolar. Os resultados mostraram que a educação da mãe e a escolaridade média do professor são determinantes da progressão individual por série e, ademais, o efeito da escolaridade do professor é substituído ao da educação materna.

Buscando explicar produtos educacionais como probabilidade de freqüentar a escola, probabilidade de freqüentar a escola no tempo adequado e progressão por série, Riani (2005) recorre, também, a aplicação de modelos hierárquicos aos dados do Censo Escolar, tendo por primeiro nível o indivíduo e, por segundo, o município. A autora conclui que, para o ensino fundamental, a idade tem relação negativa com a probabilidade de freqüentar escola, e que, esta é maior entre mulheres, brancos e amarelos, e moradores da zona urbana. Dentre as variáveis de *background* familiar (educação da mãe, domicílio chefiado por mulher, família convivente, categoria de ocupação do chefe nível superior, categoria de ocupação do chefe nível médio, chefe não ocupado), a que apresentou maior impacto sobre a probabilidade de freqüentar a escola é a educação da mãe. Para o segundo nível de análise (município), percebe que, quanto menor for o tamanho da coorte de 7 a 14 anos, maior a probabilidade de freqüentar a escola. Assim como Rios-Neto et alii (2002) verifica que o impacto da escolaridade materna é minimizado, principalmente no ensino fundamental, pelo efeito da proporção de professores de nível superior, e acrescenta o efeito de alunos por turma e de infra-estrutura. Estas três variáveis combinadas são substituídas à escolaridade materna. No caso da progressão por série, observa, por meio de

---

<sup>2</sup> Rios-Neto, E. L. G., CÉSAR, C. C., Riani, J. L. R. "Estratificação educacional e progressão escolar por série no Brasil". *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 32, n. 3, p. 395-415, Dez. 2002 *apud* Riani (2005).

simulações, que as variáveis que mais a influenciam, no primeiro e no quarto ano de estudo, são a proporção de professores com curso superior, a média de alunos por turma e infra-estrutura.

Os trabalhos mencionados mostram a importância de se considerar fatores medidos em diferentes níveis para explicar o desempenho ou algum outro produto educacional como frequência ou progressão. Variáveis referentes à família do aluno são relevantes, como também as que descrevem a escola ou mesmo os municípios ou regiões onde se localizam as escolas. Neste sentido, utilizamos um modelo hierárquico de três níveis – aluno, escola e município – para analisar como a proficiência em matemática do aluno pode ser explicada por um conjunto de fatores a ele relacionados em associação a um conjunto de características da escola que ele frequenta e do município onde reside.

### 3. Descrição do modelo hierárquico

As três fontes de variação – aluno, escola e município – podem ser incorporadas num modelo hierárquico linear em três níveis da seguinte forma:

*Nível 1:*

$$Y_{ijk} = \beta - 0jk + \beta_{1jk}X_{ijk} + \epsilon_{ijk} \quad (1)$$

onde  $Y_{ijk}$  representa a proficiência em matemática (medida em logaritmo – LNPROFIC) dos alunos de determinada escola  $j$  no município  $k$ ;  $X_{ijk}$  é uma matriz de variáveis explicativas para a proficiência dos alunos e  $\epsilon_{ijk}$  é um vetor de termos de erro *i.i.d.* que incorpora os fatores não observados que afetam a proficiência dos alunos da escola  $j$  no município  $k$ . O intercepto,  $\beta_{0jk}$  corresponde à média da proficiência dos alunos subtraída do efeito médio das variáveis  $\mathbf{X}$ , e juntamente com o vetor de coeficientes das variáveis explicativas,  $\beta_{1jk}$ , pode variar entre os indivíduos segundo a escola que eles frequentam e a sua localização. Os modelos de nível 2 e 3 levam em consideração essa ausência de independência das observações. Entretanto, embora as inclinações,  $\beta_{1jk}$ , possam diferir segundo a escola frequentada, neste trabalho consideramos que somente que o intercepto do nível 1,  $\beta_{0jk}$ , varia aleatoriamente segundo escolas e municípios.

*Nível 2:*

$$\beta_{0jk} = \alpha_{00k} + \alpha_{01k}S_{jk} + \tau_{0jk} \quad (2)$$

Onde  $\alpha_{00k}$  é a média das proficiências da escola  $j$  subtraída do efeito médio das características da escola. A matriz  $S_{jk}$  representa os atributos das escolas que os alunos frequentam, que depende das variáveis de município, representadas no nível 3.

Exemplos de atributos são: a infra-estrutura da escola, a proporção de professores com ensino superior, o número de alunos por turma, a distorção idade-série da escola, etc. O termo  $\tau_{0jk}$  representa o efeito aleatório do nível escola.

*Nível 3:*

$$\alpha_{00k} = \gamma_{000} + \gamma_{001}m_k + \mu_{00k} \quad (3)$$

Onde  $\gamma_{000}$  corresponde à média da proficiência no município  $k$ ;  $\gamma_{001}$  mede os efeitos de  $M_k$  que é a matriz de variáveis de município como PIB per capita municipal, indicadores de oferta de educação e tamanho da população em idade escolar.

Assumindo-se que os termos de erro dos três níveis são independentes e seguem uma distribuição normal com média zero e variâncias  $\sigma_\epsilon$ ,  $\sigma_\tau$  e  $\sigma_\mu$ , e substituindo-se (3) e (2) em (1) temos:

$$Y_{ijk} = \gamma_{000} + \gamma_{001}M_k + \alpha_{01k}S_{jk} + \beta_{1jk} + X_{ijk} + v_{ijk} \quad (4)$$

Onde  $v_{ijk} = \mu_{00k} + \tau_{0jk} + \epsilon_{ijk}$  é um termo de erro composto, cuja variância incorpora a variabilidade remanescente, não explicada, após a inclusão das variáveis dos três níveis. O valor esperado e a variância da proficiência em matemática dos alunos podem ser calculados como:

$$\begin{aligned} E[Y_{ijk}] &= E[\gamma_{000} + \gamma_{001}M_k + \alpha_{01k}S_{jk} + \beta_{1jk}X_{ijk} + v_{ijk}] \\ &= \gamma_{000} + \gamma_{001} * E[M_k] + \alpha_{01k} * E[S_{jk}] + \beta_{1jk} * E[X_{ijk}] \end{aligned} \quad (5)$$

$$\begin{aligned} Var[Y_{ijk}] &= Var[\gamma_{000} + \gamma_{001}M_k + \alpha_{01k}S_{jk} + \beta_{1jk}X_{ijk} \\ &\quad + v_{ijk}] = \gamma_{001}^2 * var[M_k] + \alpha_{01k}^2 * var[S_{jk}] + \beta_{1jk}^2 * var[X_{ijk}] + \sigma_\epsilon \\ &\quad + \sigma_\tau + \sigma_\mu \end{aligned} \quad (6)$$

$$Var[Y_{ijk}] = \sigma_m + \sigma_s + \sigma_x + \sigma_v \quad (7)$$

$$\sigma_m = \gamma_{001}^2 * var[M_k]; \sigma_s = \alpha_{01k}^2 * var[S_{jk}]; \sigma_x = \beta_{1jk}^2 * var[X_{ijk}] e \sigma_v = \sigma_\epsilon + \sigma_\tau + \sigma_\mu$$

Por intermédio de (6) pode-se calcular a parte da variância na proficiência atribuída às variáveis dos níveis aluno e escola, bem como a importância relativa de cada uma. É possível também estimar a variabilidade da parte remanescente representada por  $\sigma_\epsilon + \sigma_\tau + \sigma_\mu$ .

A estrutura hierárquica em três níveis pode ser bastante complexa, razão pela qual começaremos pela estimação de um modelo mais simples, que exclui a influência das variáveis explicativas. Esse modelo é chamado na literatura de modelo nulo ou incondicional Bryk e Raudenbush (2002). O objetivo dessa etapa é verificar a qual ponto a parcela da variância alocada a cada nível é significativa.

O procedimento adotado neste trabalho para calcular a parte da proficiência atribuída ao diferentes níveis foi o seguinte:

- a) em primeiro lugar, realizamos uma análise de variância com efeitos aleatórios, com a decomposição da variância entre os três níveis, excluindo as variáveis explicativas (modelo incondicional), como a seguir:

$$YB_{ijk} = \beta_{0jk} + \epsilon_{ijk} \quad (8)$$

$$\beta_{0jk} = \alpha_{00k} + \tau_{0jk} \quad (9)$$

$$\alpha_{00k} = \gamma_{000} + \mu_{00k} \quad (10)$$

b) em segundo lugar, realizamos uma análise de variância com efeitos aleatórios, com a decomposição da variância entre os três níveis, incluindo as variáveis explicativas (modelo completo). A Tabela 8 mostra o percentual da variabilidade explicada em cada nível, considerando a diferença entre a variância do modelo nulo e a variância do modelo completo.

A Equação (4) não é um modelo linear típico, tornando a estimação por MQO não apropriada. Os termos de erro aleatórios,  $v_{ijk}$ , assumem agora uma forma heterocedástica, além disso, não são independentes, uma vez que a covariância entre seus termos não é nula para alunos pertencentes à mesma escola. Deste modo, os modelos deste estudo foram estimados pelo método da máxima verossimilhança, no software HLMTP<sup>3</sup> PT.

#### 4. Fontes de Dados e Tratamento das Variáveis

As quatro fontes de dados abrangidas nesta estimação são: Sistema Mineiro de Avaliação da Educação Pública (SIMAVE), Sistema Informacional Custo Aluno (SICA), Censo Escolar e o Índice Mineiro de Responsabilidade Social (IMRS) – para a explicação do resultado educacional nas escolas públicas estaduais de Minas Gerais no ano de 2003.

O SIMAVE consiste em um método de análise do sistema de ensino realizado com base no PROEB (Programa de Avaliação da Rede Pública da Educação Básica), que aplica testes aos alunos da 4<sup>a</sup> e 8<sup>a</sup> série do ensino fundamental e 3<sup>o</sup> ano do ensino médio das escolas públicas mineiras, bem como questionários destinados à comunidade escolar. O SIMAVE adota a Teoria de Resposta ao Item (TRI) que se trata de um método de interpretação das habilidades dos alunos por item específico de um caderno de questões. Cada item corresponde a uma habilidade particular da matéria, o que permite obter o resultado total das habilidades como o produto das probabilidades de sucesso em cada item. A habilidade é dada no eixo dos números reais e, ao contrário dos exames convencionais, não possui limite máximo nem mínimo. Como recorre a escalas, pode-se comparar o desempenho entre alunos de de séries diferentes em determinada disciplina.

As variáveis do SIMAVE são utilizadas no primeiro nível de análise do modelo hierárquico, o aluno. Do SIMAVE, portanto, extraímos:

Profic: proficiência do aluno em matemática (medida pela nota dos testes);

Cor: 1 – brancos e amarelos, 0 – pretos e pardos;

Esc\_resp\_fem: 1 – responsável feminina tem 8<sup>a</sup> série completa ou mais, 0 – caso

---

<sup>3</sup> Mais detalhes sobre modelos hierárquicos podem ser encontrados em Bryk e Raudenbush (2002) e Raudenbush et alii (2000).

contrário;

Livros: 1 – mais do que 20 livros em casa, 0 – caso contrário;

Ler\_livro: 1 – lê livros, 0 – caso contrário;

Ler\_jornais: 1 – lê jornais, 0 – caso contrário;

Não Trab\_fora: 1 – não trabalha fora de casa, 0 – caso contrário;

Não Trab\_domestico: não trabalha em casa, 0 caso contrário;

Devermat: 1 – sempre faz o dever de casa de matemática, 0 – caso contrário;

Nreprovado: 1 – aluno que nunca foi reprovado, 0 – caso contrário;

Manhã: 1 – estuda no horário da manhã, 0 – caso contrário; e

Noite: 1 – estuda no horário noturno, 0 – caso contrário.

O ambiente familiar, importante nesse tipo de análise, é representado pela escolaridade da mãe (*esc\_resp\_fem*), pela existência de livros no domicílio (*livros*), pela leitura de livros e jornais (*ler\_livro* e *ler\_jornais*, respectivamente), por não trabalhar em tarefas domésticas e fora de casa, respectivamente, as variáveis denominadas *Não trab\_fora* e *Não Trab\_domestico*. As duas últimas refletindo, essencialmente, a condição socioeconômica da família do estudante.

O segundo nível de análise (*escola*) é composto por variáveis referentes ao SICA e ao Censo Escolar. O Sistema Informacional Custo Aluno – SICA – foi desenvolvido pela Superintendência de Planejamento da Secretaria de Estado de Educação de Minas Gerais em 1997. Trata-se de uma metodologia que permite o cálculo do gasto por aluno nas escolas estaduais de Minas Gerais.

As informações contempladas no SICA são originadas de outros sistemas periféricos de responsabilidade da própria Secretaria de Estado da Educação, Secretaria de Estado de Recursos Humanos e Administração e Secretaria de Estado da Fazenda. Não há registro, portanto, de receitas provenientes do setor privado como receitas de doações, de eventos organizados pela escola, de aluguel de espaço para publicidade de empresas, entre outras.

O SICA consolida as informações sobre o financiamento dos níveis de ensino<sup>4</sup> Educação Infantil, Ensino Fundamental e Ensino Médio, nas modalidades Regular, Especial e Supleância, com os respectivos componentes de custos. A metodologia contempla os custos dentro e fora da escola. Os gastos relacionados diretamente à escola são classificados como Custo Dentro e os gastos realizados com a atividade de educação, mas não efetuados com pessoal e outras despesas das escolas, são classificados como Custo Fora. No caso do Custo Dentro, os gastos com pessoal do magistério e convênios são definidos como custo direto e os gastos com pessoal administrativo e despesas gerais como indireto. No Custo Fora, todos os gastos com pessoal administrativo, convênios e despesas gerais da unidade central e das regionais da Secretária da Educação, além de aposentados e pensionistas, são considerados. Neste trabalho, utilizamos apenas o custo dentro da escola, visto que se enquadra melhor à análise, já que se refere a gastos diretamente ligados à escola – nossa unidade de observação.

---

<sup>4</sup> O banco compreende as escolas da área urbana e rural do estado de Minas Gerais.

As demais variáveis da escola são retiradas do Censo Escolar. Esta base de dados é gerenciada pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira) do Ministério da Educação. O questionário do Censo Escolar é respondido pelo diretor ou responsável de cada escola e engloba os níveis: Educação Infantil, Ensino Fundamental e Ensino Médio, nas modalidades: ensino regular, educação especial, educação de jovens e adultos e educação profissional de nível técnico. Contudo, este trabalho contempla apenas os níveis Fundamental (1º e 2º ciclos) e Médio, na modalidade de ensino regular. Desta base de dados extraíram-se as seguintes variáveis:

FNDE: 1 – escolas que recebem o Fundo Nacional de Desenvolvimento da Educação, 0 – caso contrário;

Urbana: 1 – se a escola se localiza na área urbana; 0 – caso contrário;

Laboratório: *dummy* gerada com base na combinação das três informações: possui laboratório de informática, possui laboratório de ciências e possui outros laboratórios/oficinas;

Propfprof: Proporção de professores com ensino superior na escola em relação ao total de professores das seguintes categorias:

- (i) Com Fundamental (1º grau) Incompleto;
- (ii) Com Fundamental (1º grau) Completo;
- (iii) Com Médio (2º grau) Magistério Completo;
- (iv) Com Médio (2º grau) Outra Formação Completa;
- (v) Superior (3º grau) Licenciatura Completa;
- (vi) Superior (3º grau) Completo sem Licenciatura Com Magistério; e
- (vii) Superior (3º grau) Completo sem Licenciatura Sem Magistério.

Coleta: 1 – se há coleta de lixo;

Outra importante variável do Censo inserida na análise foi o número de matrículas por turma. Esta tem a função de captar o impacto que o tamanho da turma pode causar sobre o rendimento estudantil. Neste sentido, a literatura apresenta idéias ainda não conclusivas. É o caso de Hanushek (1998) que, em uma análise descritiva, na qual utiliza dados da pesquisa Student/Teacher Achievement Ratio (STAR) – Tennessee/EUA, conclui que turmas pequenas apresentam impacto positivo sobre o desempenho escolar apenas no jardim de infância, não demonstrando qualquer efeito nos anos escolares seguintes. Já Krueger (1999), também empregando dados da STAR juntamente com variáveis como sexo, cor dos estudantes e escolaridade dos professores, encontra que a maior vantagem no atendimento em classes menores acontece no primeiro ano escolar, sendo que o impacto diminui nas séries seguintes, mas continua significativo.

No terceiro nível de análise utiliza-se a base de dados do Índice Mineiro de Responsabilidade Social. O IMRS faz parte de um projeto de criação e gerenciamento de uma ampla base de dados mineira – o DataGerais – elaborada pela Fundação João Pinheiro e coordenada pela Secretaria Estadual de Planejamento e Gestão e pela Secretaria Estadual de Desenvolvimento Econômico. O IMRS procura exibir a realidade sócio-econômica e cultural dos municípios mineiros, oferecendo uma gama de informações nas dimensões: educação, saúde,

segurança pública, emprego e renda, demografia, gestão, habitação, infra-estrutura e meio ambiente, cultura, lazer e desporto.

Inseriu-se na estimação a variável “Produto Interno Bruto per capita municipal a preços de mercado”, extraída do módulo emprego e renda, por ser o PIB per capita municipal uma boa medida do grau de desenvolvimento socioeconômico do município. Além disso, considerou-se os efeitos espaciais por meio de *dummies* regionais. São 12 mesorregiões em Minas Gerais: Noroeste, Norte, Jequitinhonha (meso de referência), Vale do Mucuri, Triângulo Mineiro/Alto Paranaíba, Central Mineira, Metropolitana de Belo Horizonte, Vale do Rio Doce, Oeste de Minas, Sul/Sudoeste de Minas, Campo da Vertentes, Zona da Mata.

Há uma considerável perda de informação referente às questões não respondidas nas três bases de dados. Das três bases, é a do SIMAVE, que conta com maior número de *missing*, porque muitos alunos deixam de responder ao questionário. Foram feitos testes de média e, realmente, há um viés de seleção, porque os alunos de pior desempenho são justamente aqueles que não responderam. Para contornar esse problema, recorreu-se ao método *hotdeck*,<sup>5</sup> com o qual imputam-se valores aos *missing* conforme as características dos “vizinhos” mais próximos. Calculou-se a média de repostas por cada variável. Mantiveram-se apenas as escolas onde a média de repostas dos alunos era igual ou acima da média geral de repostas, descartando-se aquelas onde a incidência de *missing* era superior. Em uma outra etapa, ordenou-se o banco por escolas e município, atribuindo-se valores das linhas (alunos) imediatamente mais próximos, controlados pela escola. A hipótese subjacente a este método é a de que alunos que frequentam a mesma escola apresentam, geralmente, características socioeconômicas diferentes, o que, em termos da média, não afeta os resultados obtidos.

Mesmo com a imputação de *missing*, houve redução do número de observações. A tabela seguinte evidencia tal redução:

## 5. Resultados

### 5.1. Análise descritiva das variáveis

Antes de apresentarmos os resultados da aplicação do modelo hierárquico, descrevemos as variáveis selecionadas para explicar o desempenho em matemática nas escolas públicas estaduais mineiras. Iniciando pela variável que afere o desempenho do aluno, ou seja, a proficiência média em Matemática (TAB. 2), nota-se que a média é maior no 3º ano (275,47), seguida pela 8ª série (243,42) e pelo 4º ano do ensino fundamental (191,12). É natural que os alunos do 3ª série

<sup>5</sup> Para maiores detalhes, ver Barzi & Woodward (2004) em “Imputations of Missing Values in Practice: Results from Imputations of Serum Cholesterol in 28 Cohort Studies”. *American Journal of Epidemiology* 160(1):34-45, July 1, 2004 e Clayton & Mander (1999) em *Statistical Software Components S*, 1999 – ideas.repec.org.

Tabela 1

Número de observações nas bases de dados segundo nível de análise e etapa educacional

Nível de análise	Série	Total de obs.	Obs.sem <i>missing</i>	Obs.depois <i>hotdeck</i>
	4a	140667	94614	96174
Primeiro nível:	8a	195333	91109	137913
aluno	3a	117711	72941	81737
Segundo nível:	4a	2469	1635	1833
escola	8a	2552	1949	1761
	3a	1453	1404	1039

Fonte: Simave e Censo Escolar, 2003.

apresentem um desempenho médio melhor do que o das demais séries, uma vez que devem acumular habilidade ao longo do tempo.<sup>6</sup>

Tabela 2

Proficiência média dos alunos segundo nível de ensino – 2003

Série	Proficiência média	Desvio padrão
4a fundamental	191,12	48,66
8a fundamental	243,42	49,29
3o médio	275,47	50,58

Fonte: Simave, 2003.

As variáveis referentes ao primeiro nível de análise (aluno) mostram que não há muita distinção entre as características dos alunos da 4ª série, 8ª série do Ensino Fundamental e 3ª série do Ensino Médio (TAB. 3). À exceção dessa última série, a proporção de brancos e amarelos é inferior à de negros e pardos. Como se trata da rede pública, tais grupos demográficos devem prevalecer, uma vez que estão sobre-representados entre os mais pobres, sendo provenientes de famílias sem recursos para arcar com a educação na rede privada.

A escolaridade da mãe é baixa, uma vez que cerca de 70% não tem a 8ª série completa. Embora apenas entre 20 e 30% dos domicílios de alunos tenha mais de 20 livros, acima de 70% dos alunos lêem livros e isto é tanto mais evidenciado quanto menor o nível de escolaridade. No caso da leitura de jornais, ocorre o inverso, aproximadamente 65% dos alunos do ensino médio realizam este tipo de leitura.

Menos de um terço dos alunos nas três séries abordadas neste trabalho realiza tarefas domésticas. No caso de participação no mercado de trabalho, a proporção de alunos da 4ª série que trabalham é de 14,36%, chegando, contudo, a 50,60% no 3º ano.

<sup>6</sup> O método TRI, ao estabelecer a aplicação de questões comuns nas provas, possibilita avaliar o ganho de habilidade ao longo do processo de formação educacional. Em média, esse ganho é observado, porém um aluno da 4ª série do Ensino Fundamental pode apresentar desempenho melhor do que o esperado para o seu nível de habilidade em relação ao da 3ª série do Ensino Médio, mesmo que o nível de proficiência do aluno mais velho venha a ser superior ao da 4ª série.

Na 4ª série, mais de 70% fazem o dever de Matemática e essa proporção cai até o 3º ano, 40,79%. Uma justificativa real, embora negativa, para esta redução é a maior proporção de alunos do ensino médio trabalhando. A parcela de não reprovados é, também, menor para os alunos do 3º ano, cerca de 64%.

Os alunos da 4ª e 8ª séries estudam, em maior parcela, no turno da manhã e os de 3º ano à noite, muitos devido à necessidade de compatibilizar estudo e trabalho.

Tabela 3

Variáveis dos alunos: proporções segundo nível de ensino – 2003

	4a série EF	8a série EF	3a série EM
Branços	41,47	48,23	51,20
Respons.c/pelo menos 8a série completa	39,38	23,67	25,47
Possuem mais de 20 livros em casa	20,07	27,70	30,04
Lêem livros	80,15	71,03	70,90
Lêem jornais	47,03	47,41	65,36
Não trabalham fora	85,8	66,45	49,75
Não trabalham em casa	69,65	72,45	72,79
Fazem dever de matematica	73,75	54,72	40,79
Nunca foram reprovados	77,44	69,38	64,52
Estudam de manhã	59,22	63,36	42,67
Estudam à tarde	40,60	23,00	2,79
Estudam à noite	0,19	13,64	54,56

Fonte: Simave, 2003.

No que tange ao segundo nível, constatamos que mais de 70% recebem recursos do FNDE e esta proporção é maior quanto mais introdutória é a série. Quanto mais avançada a série, maiores as chances da escola se localizar na área urbana (URB). A presença de laboratório (Laborato) é menor nas escolas que ofertam a 4ª série (29,57%) e maior nas de 3ª série do ensino médio (45,57%). A presença de coleta de lixo (Coleta), variável indicadora do acesso a infra-estrutura urbana, é elevada, posto que mais de 80% das escolas são providas deste serviço.

Tabela 4

Variáveis da escola: proporções segundo nível de ensino – 2003

	4 <sup>a</sup>	8 <sup>a</sup>	3 <sup>a</sup>
Variáveis	Fundamental	Fundamental	Médio
FNDE	86,42	81,77	79,09
URB	84,72	89,84	98,99
Laborato	29,57	42,02	45,57
Coleta	96,56	87,56	92,82

Fonte: Censo Escolar, 2003.

O valor do custo dentro médio mensal por aluno é praticamente o mesmo para as três séries (R\$ 62,00), o que distingue é o desvio padrão da 3<sup>a</sup> série do Ensino Médio bem superior aos das demais séries.

A proporção de professores com ensino superior é maior nas escolas de 8<sup>a</sup> série (81%) e bem menor nas de 4<sup>a</sup> série (44%), de acordo com a TAB. 5. Chama atenção a menor proporção de professores com tal formação na 3<sup>a</sup> série do médio (49%) relativamente a 8<sup>a</sup> do fundamental. Entretanto, o desvio padrão é muito próximo ao valor da média, sugerindo maior heterogeneidade entre as escolas que ofertam este nível. As turmas da última série do Ensino Médio são maiores do que as da 4<sup>a</sup> série do Ensino Fundamental e apresentam tamanho bem próximo as de 8<sup>a</sup> série, algo esperado nesta etapa do ciclo educacional.

Tabela 5

Variáveis da escola: Valores médios segundo nível de ensino – 2003

	4 <sup>a</sup>	8 <sup>a</sup>	3 <sup>a</sup>
	Fundamental	Fundamental	Médio
CDENTRO*	62,62 (72,17)	61,04 (86,47)	62,03 (110,55)
PROPPROF	0,44 (0,26)	0,81 (0,21)	0,49 (0,43)
MATTURMA*	28,79 (6)	33,23 (7,05)	34,22 (7,22)

(\*)Variáveis medidas com base na média mensal do ano por escola. O custo dentro foi corrigido pelo IPC-m(FGV).

Desvios padrões entre parentêsis.

Fonte: SICA (2003) e Censo Escolar (2003).

Observando as variáveis dos municípios, percebe-se que o maior PIB per capita ocorre nos municípios dotados da última série do Ensino Médio em relação às outras séries analisadas (TAB. 6). O resultado não é surpreendente, porque estes municípios “mais ricos” devem contar com área urbana expressiva onde se concentram as escolas públicas que ofertam o Ensino Médio, a relação é, portanto, a esperada. Por fim, há maior concentração de alunos na meso de Mucuri, Zona da Mata, Sul, Central Mineira e Oeste de Minas.

Tabela 6

Variáveis do município: valores médios de PIB per capita municipal e proporção de alunos por mesoregião segundo nível de ensino – 2003

Variáveis	4 <sup>a</sup> Fundamental	8 <sup>a</sup> Fundamental	3 <sup>a</sup> Médio
PIBPC média	4745,97	5017,57	5293,41
desvio	3260,44	4427,06	6886,89
Campo das Vertentes	2,99	2,69	2,18
Central Mineira	13,15	12,16	11,23
Jequitinhonha	9,16	7,27	7,02
Metropolitana	4,38	2,69	2,96
Noroeste	6,97	7,42	7,18
Norte	3,78	3,16	4,06
Oeste	10,56	9	13,42
Sul	12,95	13,27	11,54
Triângulo Mineiro	3,98	4,9	5,62
Mucuri	14,74	18,01	16,38
Rio Doce	3,19	4,58	4,99
Zona da Mata	14,14	14,85	13,42

PIBPC – Produto Interno Bruto a preços de mercado.

\*\*Valores atualizados em R\$ de 2004 pelo IGP-DI.

Fonte: Elaboração própria a partir das bases de dados do Censo Escolar/2003 e do SICA/2003.

## 5.2. Resultados dos modelos hierárquicos.

### 5.2.1. Modelo nulo – Decomposição da variância

A Tabela 7 traz, na última coluna, o percentual da variabilidade explicada em cada nível, considerando a diferença entre a variância do modelo nulo e a variância do modelo completo.

A proporção da variância explicada pelos três níveis difere segundo as séries no caso dos níveis escola e município. Em relação ao primeiro nível, a variabilidade da proficiência que é devida aos atributos dos alunos corresponde à cerca de 37% da variabilidade total na 4<sup>a</sup> série, à 57% na 8<sup>a</sup> série e 67% na 3<sup>a</sup> série do Ensino Médio, ou seja, à medida em que avançamos no ciclo educacional, a heterogeneidade entre os alunos tende a explicar mais o seu desempenho.

Para a quarta e oitava série, a importância das características da escola é praticamente a mesma, na casa dos 20%. Na última série do Ensino Médio,

a variabilidade das escolas é maior, aproximadamente 45%, sendo que as características dos municípios contribuem negativamente com  $-12,45\%$ . Este resultado negativo pode ocorrer por diversas razões, entre elas, erro amostral, problemas de especificação do modelo, ou mesmo problemas de performance nos algoritmos usados por alguns softwares.<sup>7</sup> A razão que poderia ser apontada como mais provável no caso deste estudo é a concentração de escolas públicas estaduais que ofertam ensino médio em municípios maiores, onde não há muita diferenciação. A alternativa seria fixar variância como zero ou retirar o terceiro nível (municípios) do modelo. Entretanto, optamos por não realizar este procedimento, porque além de não mudar em nada os resultados para os outros níveis, o nosso objetivo é a comparação das diferentes séries de ensino com os mesmo modelos.

Tabela 7

Variáveis dos alunos: proporções segundo nível de ensino – 2003

4a série do Ensino Fundamental				
	Modelo nulo	Modelo com os 3 níveis	Diferença	Var explicada (%)
Partição da variância				
Município	0.006	0.00156	0.00444	36.54
Escola	0.0097	0.0065	0.0032	26.34
Aluno	0.0527	0.04819	0.00451	37.12
Total	0.068	0.05625	0.01215	100
8a série do Ensino Fundamental				
	Modelo nulo	Modelo com os 3 níveis	Diferença	Var explicada (%)
Partição da variância				
Município	0.00254	0.00116	0.00138	19.03
Escola	0.00431	0.00257	0.00174	24
Aluno	0.03779	0.03366	0.00413	56.97
Total	0.04464	0.03739	0.00725	100
3a série do Ensino Médio				
	Modelo nulo	Modelo com os 3 níveis	Diferença	Var explicada (%)
Partição da variância				
Município	0.00088	0.00152	-0.00064	-12.45
Escola	0.0044	0.00208	0.00232	45.14
Aluno	0.02982	0.02636	0.00346	67.31
Total	0.0351	0.02996	0.00514	100

Fonte: Elaboração própria a partir das bases de dados do Censo Escolar/2003, do SIMAVE/2003 e do SICA/2003.

<sup>7</sup> Para maiores detalhes, ver *Performance Assessment for the Workplace, Volume II: Technical Issues* Alexandra K. Wigdor and Bert F. Green, Jr., Editors; Committee on the Performance of Military Personnel, National Research Council.

Os resultados dos modelos hierárquicos para a 4<sup>a</sup> e 8<sup>a</sup> séries do fundamental e 3<sup>a</sup> série do ensino médio estão apresentados nas Tabelas 8, 9 e 10. Como se observa, a maioria das variáveis é significativa para o nível do aluno.

Sobre a estimação para a 4<sup>a</sup> série (TAB. 8), à exceção de turno da noite e cor, quase todas as variáveis de nível 1 são significativas (e positivas) para explicar a proficiência em matemática dos alunos. As variáveis, por ordem de importância, são:

1. Nunca ter sido reprovado (NREPROV) aumenta em cerca de 10%<sup>8</sup> o desempenho em Matemática;
2. Ler livros (LELIVRO) melhora em 5,66 % a proficiência de alunos de 4<sup>a</sup> série;
3. Não trabalhar fora (NTRABFORA), em 4,97%;
4. Fazer o dever de matemática (FDEVMAT) melhora em, aproximadamente, 4,66% o rendimento na disciplina;
5. A responsável feminina com 8<sup>a</sup> série completa ou mais (ESCMAE) aumenta em 4,32% a proficiência em matemática;
6. Ler jornais (LEJORNAIS), em 3,33%.

No nível das escolas, destacam-se presença de coleta de lixo (COLETA) com sinal negativo (-3,41%) e o fato de se localizar na área urbana (2,54%). O primeiro resultado é contra-intuitivo pois, a princípio espera-se que uma escola provida de serviços urbanos deva contribuir em maior medida para a boa qualidade de ensino. Além disso, o resultado da variável “urbana” (URB) é mais um indício dessa relação positiva. Como mostra a Tabela 4, 96% das escolas da 4<sup>a</sup> série apresentam coleta de lixo, assim, como quase todas tem acesso a sua influência.

Em termos de municípios, o fato dos alunos residirem nas mesos do Triângulo Mineiro, Vale do Rio Doce e Mucuri contribui com 10%, 9% e 8% para o desempenho em Matemática, ao passo que na Central Mineira e Região Metropolitana de Belo Horizonte o efeito é negativo, em torno de 5%. Quanto mais rica a região, medida pelo logaritmo do PIB per capita (LNPIBPC), melhor o desempenho dos alunos. Trata-se, na verdade, de uma medida de condições socioeconômicas favoráveis do município atuando positivamente sobre a qualidade do ensino, em especial, em matemática.

<sup>8</sup> Para calcularmos o efeito sobre a variável dependente, utilizamos o antilog dado por  $[exp(\beta) - 1]x100$ .

Tabela 8

## Resultados dos modelos – 4ª série do Ensino Fundamental

	Modelo nulo		Modelo 1		Modelo 2		Modelo 3	
	Valor	t	Valor	t	Valor	t	Valor	t
NIVEL 3: 502 obs								
INTERCEPTO	5.191	1093.43	4.9659	647.27	4.979945	336.5	4.952015	282.778
LNPIBPC							0.017482	2.555
CAMPO							0.017724	0.828
CENTRAL							-0.059529	-3.875
METROBH							-0.054208	-3.606
NORO							0.03839	2.036
NORTE							0.048759	2.052
OESTE							0.053131	3.495
SUL							0.015409	1.14
TRIANG							0.096101	5.283
MUCURI							0.08443	5.767
RDOCE							0.086507	4.545
ZMATA							0.039777	2.857
NIVEL 2: 1833 obs								
FNDE					-0.00101	-0.132	-0.000811	-0.113
CDENTRO					0.000131	4.625	0.000133	4.726
URB					0.037381	4.666	0.025098	3.217
LABORATO					0.008003	1.657	0.005096	1.096
COLETA					-0.04943	-3.00	-0.034752	-2.276
PROPPROF					0.023034	1.777	0.023204	1.811
MATTURMA					0.001527	2.881	0.001689	3.18
NIVEL 1: 96174 obs								
MANHA			0.0111	1.83	0.011551	1.913	0.011994	2.01
NOITE			0.0677	0.87	0.065781	0.843	0.064013	0.804
COR			0.0004	0.19	0.000407	0.197	0.000109	0.053
ESCMAE			0.0427	20.30	0.042338	20.1	0.042252	20.042
TEMLIVRO			0.0343	17.11	0.01706	17.02	0.016967	16.881
LELIVRO			0.0550	24.57	0.055046	24.55	0.055015	24.529
LEJORNALIS			0.0331	19.33	0.032985	19.29	0.032768	19.175
NTRABFORA			0.0488	16.98	0.048701	16.97	0.048537	16.893
NTRABCASA			-0.0333	-18.66	-0.03327	-18.67	-0.033241	-18.666
FDEVMAT			0.0456	21.38	0.045632	21.45	0.04559	21.366
NREPROV			0.0957	33.36	0.095526	33.35	0.09526	33.278
Partição da Variância								
Município	0.006***		0.00429***		0.00406		0.00156	
Escola	0.0097***		0.00691***		0.0066		0.0065	
Aluno	0.0527***		0.04819***		0.0482		0.04819	

Fonte: Elaboração própria a partir das bases de dados do Censo Escolar/2003, do Simave/2003 e do SICA/2003

No caso dos resultados da 8<sup>a</sup> série (TAB. 9), à exceção de turno da manhã, todas as variáveis de nível 1 são significativas para explicar a proficiência em matemática dos alunos. As variáveis, por ordem de importância, são:

1. Assim como na 4<sup>a</sup> série, nunca ter sido reprovado (NREPROV) aumenta em 10% o desempenho;
2. A escolaridade da mãe superior a 8<sup>a</sup> série (ESMAE), em 4%;
3. Estudar no turno da noite (NOITE) reduz em 3%;
4. Fazer dever de matemática (FDEVMT) melhora o desempenho nesta disciplina em 3%;
5. Ter mais de 20 livros em casa (TEMLIVRO) amplia em 3%.

Em termos do nível escola, todas as variáveis não são significativas, apenas custo tem efeito, embora bastante inexpressivo 0,012%. No que tange ao 3<sup>o</sup> nível, municípios, o fato dos alunos residirem nas mesos do Triângulo Mineiro e Mucuri contribuem em 5% para o desempenho em Matemática. Mais uma vez, residir e estudar na Central Mineira e Região Metropolitana de Belo Horizonte apresenta efeito negativo sobre a proficiência de cerca de 6%.

A Tabela 10 apresenta os resultados para a 3<sup>a</sup> série do ensino médio. Novamente, as variáveis referentes a alunos apresentam maior poder de explicação e são semelhantes às das demais séries.

1. Nunca ter sido reprovado (NREPROV) melhora em 7% o desempenho;
2. Estudar no turno da manhã aumenta em 2,6%;
3. A escolaridade da mãe (ESMAE), em 2,5%;
4. Possuir mais de vinte livros (TEMLIVRO) em casa melhora em 2% o desempenho;

Por outro lado, não contribuir para os afazeres domésticos (NTRABCASA) e estudar à noite reduzem a proficiência em 5% e 3%, respectivamente. Possivelmente, essas duas variáveis estejam mascarando o efeito negativo de trabalhar fora de casa sobre o desempenho educacional, uma vez que se o aluno não ajuda em casa e estuda à noite é maior a chance de já ter ingressado no mercado de trabalho, o que reduziria o tempo alocado para o estudo, afetando negativamente a proficiência em matemática. No que tange a escola, a presença de coleta de lixo (COLETA) aumenta em 2,7% a proficiência e de laboratórios (LABORATO), em 1,07%. Esta última variável tem importância, na medida em que os alunos do ensino médio são introduzidos às disciplinas de Física e Química e o ensino de qualidade das mesmas depende em grande medida de laboratórios. Se a escola é provida desses equipamentos, a qualidade do ensino deve ser melhor nas áreas exatas, o que envolve a matemática. No terceiro nível, municípios, nenhuma das variáveis é estatisticamente significativa.

Tabela 9

## Resultados dos modelos – 8ª série do Ensino Fundamental

NIVEL 3: 633 obs	Modelo nulo		Modelo 1		Modelo 2		Modelo 3	
	Valor	est_t	Valor	est_t	Valor	est_t	Valor	est_t
INTERCEPTO	5.457	1867.10	5.310	651.30	5.286875	597.449	5.29411	431.015
LNPIBPC							0.007123	1.626
CAMPO							-0.012109	-0.725
CENTRAL							-0.062106	-5.846
METROBH							-0.058886	-4.3
NORO							0.017863	1.34
NORTE							0.02055	1.49
OESTE							0.015213	1.271
SUL							0.002887	0.258
TRIANG							0.04151	3.289
MUCURI							0.043508	3.973
RDOCE							0.023035	1.752
ZMATA							0.020841	1.767
NIVEL 2: 1761 obs								
FNDE					-0.005117	-1.245	-0.005064	-1.287
CDENTRO					0.000119	6.483	0.000119	6.408
URB					0.013608	1.914	0.00996	1.456
LABORATO					-0.001205	-0.391	-0.00148	-0.5
COLETA					0.018025	2.98	0.005033	0.854
PROPPROF					-0.022282	-1.869	-0.017911	-1.48
MATTURMA					0.00057	1.668	0.000658	1.915
NIVEL 1: 137913 obs								
MANHA			0.004305	1.217	0.004135	1.183	0.003834	1.105
NOITE			-0.03716	-9.671	-0.037265	-9.768	-0.037061	-9.791
COR			0.016713	13.377	0.016714	13.387	0.016364	13.08
ESCMAE			0.039733	21.431	0.039497	21.201	0.03953	21.195
TEMLIVRO			0.031393	24.705	0.031236	24.668	0.031147	24.602
LELIVRO			0.019137	13.7	0.019133	13.73	0.019134	13.728
LEJORNAIS			0.015247	10.301	0.015186	10.235	0.015066	10.108
NTRABCASA			0.026137	19.419	0.026044	19.353	0.025936	19.199
NTRABFORA			0.011667	5.921	0.011604	5.892	0.011608	5.901
FDEVMAT			0.034703	26.32	0.03471	26.364	0.034889	26.49
NREPROV			0.095747	69.69	0.095746	69.613	0.095611	69.441
Partição da Variância								
Município	0.00254***		0.00237***		0.00228		0.00118	
Escola	0.00431***		0.00276***		0.00259		0.00257	
Aluno	0.03779***		0.03366***		0.03366		0.03366	

Fonte: Elaboração própria a partir das bases de dados do Censo Escolar/2003, do Simave/2003 e do SICA/2003

Tabela 10

## Resultados dos modelos – 3ª série do Ensino Médio

	Modelo nulo		Modelo 1		Modelo 2		Modelo 3	
	Valor	est.t	Valor	est.t	Valor	est.t	Valor	est.t
NIVEL 3: 641 obs								
INTERCEPTO	5.594098	2052.299	5.56228	716.982	5.522831	302.076	5.514836	271.445
LNPIBPC							0.003576	0.657
CAMPO							-0.021248	-1.331
CENTRAL							0.003716	0.346
METROBH							0.018941	1.484
NORO							0.002977	0.212
NORTE							0.036405	2.058
OESTE							-0.005713	-0.524
SUL							0.014241	1.295
TRIANG							0.007546	0.536
MUCURI							0.010215	0.965
RDOCE							0.005295	0.444
ZMATA							0.008872	0.79
NIVEL 2: 1039 obs								
FNDE					-0.00954	-2.345	-0.009626	-2.347
CDENTRO					0.000102	7.562	0.000102	7.542
URB					0.016186	0.9	0.01895	1.044
LABORATO					0.010244	2.74	0.010636	2.878
COLETA					0.028539	4.1	0.026747	3.891
PROPPROF					-0.00581	-1.113	-0.005921	-1.135
MATTURMA					0.00051	0.782	0.000493	0.76
NIVEL 1: 81737 obs								
MANHA			0.02611	3.217	0.026014	3.241	0.026068	3.242
NOITE			-0.03237	-4.133	-0.03213	-4.139	-0.032044	-4.119
COR			0.012141	10.039	0.012116	10.003	0.012111	10.002
ESCMAE			0.024425	15.56	0.024314	15.531	0.024325	15.529
TEMLIVRO			0.020373	14.217	0.020282	14.181	0.02026	14.181
LELIVRO			0.002253	1.209	0.002271	1.226	0.002273	1.226
LEJORNAIS			0.021447	17.869	0.021372	17.871	0.021356	17.859
NTRABCASA			-0.04926	-31.636	-0.04915	-31.775	-0.049143	-31.796
NTRABFORA			-0.00758	-4.007	-0.00753	-3.994	-0.007523	-3.987
FDEVMAT			0.01818	4 12.926	0.018243	13.012	0.018242	13.01
NREPROV			0.07649	4 57.223	0.076483	57.305	0.076499	57.341
Partição da Variância								
Município	0.00088***		0.0015***		0.00158		0.00152	
Escola	0.0044***		0.00238***		0.0021		0.00208	
Aluno	0.02982***		0.02636***		0.02636		0.02636	

Fonte: Elaboração própria a partir das bases de dados do Censo Escolar/2003, do Simave/2003 e do SICA/2003

## 6. Considerações Finais

Os resultados deste trabalho constituem-se em mais uma evidência de que o efeito da escola e do município de residência sobre a variabilidade observada no aproveitamento do aluno é reduzido frente às características do aluno e do *background* familiar, mesmo aplicando a classe de modelos hierárquicos, que são reconhecidos por separarem os efeitos de componentes de vários níveis e, por isso, tendem a não subestimar o efeito de cada um deles sobre a qualidade do ensino.

Além do recurso ao modelo hierárquico, as variáveis da escola abrangeram uma série de aspectos tais como custo do aluno, acesso a fundos públicos, tamanho de turmas, capacitação dos professores, presença de laboratórios, infra-estrutura urbana da escola e, ainda assim, o poder de explicação dessas variáveis é relativamente baixo. Basicamente, a infra-estrutura-urbana, por meio seja de coleta de lixo seja da localização em área urbana, explica a proficiência em matemática nos resultados do modelo. Apenas na 3ª série do Ensino Médio, a presença de laboratórios distingue o desempenho em relação às demais séries. É, também, para essa etapa que a variabilidade entre as escolas conta mais na explicação da qualidade do ensino em Matemática.

Em se tratando do terceiro nível, municípios, identifica-se o maior peso para a 4ª série, onde não só o PIB per capita é importante assim como a localização das escolas em várias mesoregiões. Entretanto, para as séries mais avançadas, a variabilidade entre os municípios perde importância, chegando, no caso da 3ª série do ensino médio, a contribuir negativamente e sem que nenhuma variável no modelo apresente significância estatística.

Deste modo, em termos das características do aluno, chama atenção, em todos as séries analisadas, o efeito positivo da não-reprovação, da escolaridade da mãe e de possuir, pelo menos, um reduzido acervo de livros em casa. No caso das séries do ensino fundamental, além desses efeitos, são importantes e comuns os impactos positivos de fazer o dever de matemática em casa. Por fim, na 3ª série do ensino médio, não alocar tempo para afazeres domésticos se constitui em uma das variáveis explicativas ao contrário das outras etapas analisadas.

Tais resultados, no entanto, não minimizam o papel de políticas públicas para melhoria da educação, uma vez que há correlação positiva entre *background* familiar e qualidade das escolas. Investir em escolas em tempo integral, treinamento de professores e na modernização de equipamentos, certamente, reduzirá o efeito da família, ampliando o papel da escola.

## Referências bibliográficas

- Afonso, A. & Aubyn, M. (2005). Cross-country efficiency of secondary education provision: A semiparametric analysis with non-discretionary inputs. European Central Bank Working Paper Series 494.
- Albernaz, A., Ferreira, F. H. G., & Franco, C. (2002). Qualidade e equidade na educação fundamental brasileira. *Pesquisa e Planejamento Econômico*, 32(3):453–476.
- Barbosa, M. E. F. & Fernandes, C. (2001). A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4a. série. In Franco, C., editor, *Avaliação, Ciclos e Promoção na Educação*, pages 155–172. ArtMed, Porto Alegre.
- Barros, R. P., Mendonça, R., Santos, D. D., & Quintaes, G. (2001). Determinantes do desempenho educacional do Brasil. *Pesquisa e Planejamento Econômico*, 31(1):1–42.
- Bryk, A. S. & Raudenbush, S. W. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage, Newbury Park, California, 2nd. edition. 485 p.
- César, C. C. & Soares, J. F. (2001). Desigualdades acadêmicas induzidas pelo contexto escolar. *Revista Brasileira de Estudos Populacionais*, 18(1/2):97–110.
- Felício, F. & Fernandes, R. (2005). O efeito da escola sobre o desempenho escolar: Uma avaliação do ensino fundamental no estado de São Paulo. In *Anais Do XXXIII Encontro Nacional de Economia*. ANPEC.
- Ferrão, M. E., Beltrão, K. I., Fernandes, C., Santos, D., Suárez, M., & Andrade, A. C. (2001). O SAEB – Sistema Nacional de Avaliação da Educação Básica: Objetivos, características e contribuições na investigação da escola eficaz. *Revista Brasileira de Estudos Populacionais*, 18(1/2):111–130.
- Fletcher, P. R. (1998). *À Procura do Ensino Eficaz*. MEC, Rio de Janeiro.
- Hanushek, E. A. (1998). The evidence on class size. Rochester, N.Y.: University of Rochester, W. Allen Wallis Institute of Political Economy (Occasional Paper 98-1).
- Hanushek, E. A., Gomes-Neto, J. B., & HARBISON, R. W. (1996). Efficiency-enhancing investments in school quality. In Birdsall, N. & Sabot, R. H., editors, *Opportunity Forgone: Education in Brazil*, pages 385–424. Inter-American Development Bank, Washington, D.C.
- Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2):497–532.
- Raudenbush, S., Bryk, A. S., Cheong, Y. F., & Congdon Jr, R. T. (2000). *Hierarchical Linear and Nonlinear Modeling*. Scientific Software International, Lincolnwood.
- Riani, J. L. R. (2005). Determinantes do resultado educacional no Brasil: Família, perfil escolar dos municípios e dividendo demográfico numa abordagem hierárquica e espacial. Belo Horizonte, UFMG/Cedeplar.
- Rios-Neto, E. L. G., César, C. C., & Riani, J. L. R. (2002). Estratificação educacional e progressão escolar por série no Brasil. Artigo apresentado no Encontro Determinantes do Sucesso Educacional, IPEA.
- Silva, N. V. & Hasenbalg, C. (2001). Recursos familiares e transições educacionais. Versão preliminar apresentada no Workshop de Demografia da Educação. Associação Brasileira de Estudos Populacionais, ABEP, Salvador, Bahia.