

Qualidade do ensino em Matemática: determinantes do desempenho de alunos em escolas públicas estaduais mineiras

Ana Flávia Machado(CEDEPLAR/UFMG)

Sueli Moro (CEDEPLAR/UFMG)

Laudiemy Martins (CEDEPLAR/UFMG)

Juan Rios (CEDEPLAR/UFMG)

1 - Introdução

Os indicadores referentes à educação, no Brasil, tais como taxa de analfabetismo, taxa de atendimento escolar, taxa de escolarização líquida, taxa de evasão, qualificação dos professores 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.

Em parte, estes resultados são provenientes de políticas educacionais que têm buscado a universalização do acesso ao ensino. O exemplo mais marcante desta experiência é o FUNDEF (Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização dos Trabalhadores na Educação)¹. Por meio de transferências de recursos da União para Estados e Municípios, o FUNDEF pretendeu assegurar a entrada e a manutenção da população em idade escolar no ensino fundamental.

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? Proficiência em disciplinas-chave dos currículos certamente é um deles. 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 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 e, em um segundo nível, as características da escola, incluindo o

¹ O FUNDEB (Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Trabalhadores da Educação), em tramitação no Congresso Nacional, deve substituir o FUNDEF. Os objetivos são praticamente os mesmos do FUNDEF, com exceção da extensão ao ensino médio.

custo-aluno. Estimacões semelhantes, recorrendo a modelos hierárquicos, já foram feitas por Fletcher (1998), César e Soares (2001), Barbosa e Fernandes (2001), Ferrão *et al* (2001), Albernaz, Ferreira e Franco (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)², 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. No Brasil, são escassas as informações sobre custo-aluno.

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

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, além de insumos escolares:

$$Y = F(c_a, c_f, c_p, \varepsilon)$$

Y - desempenho dos alunos medido pela proficiência;

c_a - vetor de características dos alunos tais como cor, sexo, idade;

c_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;

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

ε – termo de erro aleatório

No âmbito de trabalhos que utilizam metodologias que não são as de modelos hierárquicos, Felício e Fernandes (2005) argumentam que, em regra, o efeito da escola é reduzido quando comparado à parcela explicada pelo *background* familiar, mas 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 é maior impacto das variáveis de *background* familiar até metade do ensino fundamental.

O trabalho de Barros *et alli* (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

² As bases SIMAVE e SICA são explicadas na quarta seção referente a fonte de dados e tratamento das variáveis.

fundamental, diminuindo no ensino médio) tem efeito tão relevante quanto, ou maior que, a escolaridade dos professores.

Por outro lado, Hanushek, Gomes-Neto e Harbison (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. Ao passo que 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, Ferreira e Franco (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. A inclusão das variáveis no segundo modelo causou uma queda da variância residual para 19.6% do seu valor no modelo inicial, implicando que 80.4% da variância entre as escolas derivam da composição social interna e da seletividade da escola.

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, César e Riani (2002)³ utilizam a abordagem de modelos hierárquicos, sendo o primeiro nível o indivíduo e o 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, sendo, ainda, o efeito da escolaridade do professor substituto ao da educação materna.

³ 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).

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. 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, César e Riani (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 substitutas à escolaridade materna. No caso da progressão por série, observa, por meio de simulações, que as variáveis que mais a influenciam, no primeiro e no quarto ano de estudo, são proporção de professores com curso superior, média de alunos por turma e fator de infra-estrutura. Nota, ainda, que o efeito da educação materna sobre a variável dependente é reduzido pelo impacto de proporção de matrículas em estabelecimentos com ciclos, proporção de professores com curso superior e infra-estrutura da escola.

Os trabalhos mencionados mostram a importância de se considerar fatores medidos em diferentes níveis, para explicar o desempenho ou algum outro produto educacional como freqüência 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 dois níveis, aluno e escola, 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 freqüenta.

3 - Descrição do modelo hierárquico

As duas fontes de variação, aluno e escola, podem ser incorporadas num modelo hierárquico linear da seguinte forma:

$$Y_i = b_0 + b_1 X_i + \varepsilon_i \quad (1)$$

onde Y_i representa a proficiência em matemática (medida em logaritmo - LNPROFIC) dos alunos de determinada escola k ; X_i é uma matriz de variáveis explicativas para a proficiência dos alunos e ε_i um vetor de termos de erro i.i.d. que incorpora os fatores não observados que afetam a proficiência dos alunos da escola.

Considere agora o mesmo modelo estimado para alunos de outra escola l . Caso os parâmetros estimados para os dois modelos sejam diferentes podemos dizer que as escolas diferem quando à proficiência dos seus alunos. Ou seja, alunos semelhantes que freqüentam escolas diferentes podem ter níveis de proficiência distintos. Isso acontece porque os alunos não estão distribuídos aleatoriamente entre as duas escolas.

Um modelo estimado para uma população inteira de escolas (e alunos) tomaria a seguinte forma:

$$Y_{ij} = b_{0j} + b_{1j} X_{ij} + r_{ij} \quad (2)$$

Onde i é o conjunto de alunos e j as escolas que eles freqüentam. O índice j nos parâmetros significa que cada escola agora tem seus próprios parâmetros, e o termo r_{ij} é o efeito aleatório do nível 1. A proficiência de cada escola seria, portanto, representada respectivamente por b_{0j} e b_{1j} .

Embora as inclinações, b_{1j} , possam diferir segundo a escola freqüentada, neste trabalho consideramos que somente o intercepto do nível 1, o qual representa a média da proficiência subtraída do efeito médio das outras variáveis, varia aleatoriamente segundo as escolas. Neste caso, a variável resposta no nível 2 seria:

$$b_{0j} = \lambda_{00} + \lambda_{01} S_j + u_{0j} \quad (3)$$

Onde a matriz S_j representa os atributos das escolas que os alunos freqüentam.

Exemplos de atributos são: a infra-estrutura da escola, a proporção de professores com ensino superior, o número de alunos por turma, a proficiência média da escola, etc. O termo u_{0j} representa o efeito aleatório do nível escola. Assumindo-se que os termos de erro dos dois níveis são independentes e seguem uma distribuição normal com média zero e variâncias σ_r e σ_u e substituindo-se (3) em (2) temos:

$$Y_{ij} = \lambda_{00} + \lambda_{01} S_j + b_{1j} X_{ij} + v_{ij} \quad (4)$$

Onde $v_{ij} = u_{0j} + r_{ij}$ é um termo de erro composto, cuja variância incorpora a variabilidade remanescente, não explicada, após a inclusão das variáveis dos níveis 1 e 2. O valor esperado e a variância da proficiência em matemática dos alunos após incluídas as variáveis dos dois níveis podem ser calculados como:

$$E[Y_{ij}] = E[\lambda_{00} + \lambda_{01} S_j + b_{1j} X_{ij} + v_{ij}] = \lambda_{00} + \lambda_{01} * E[S_j] + b_{1j} * E[X_{ij}] \quad (5)$$

$$Var[Y_{ij}] = Var[\lambda_{00} + \lambda_{01} S_j + b_{1j} X_{ij} + v_{ij}] = \lambda_{01}^2 * var[S_j] + b_{1j}^2 * var[X_{ij}] + \sigma_r + \sigma_u. \quad (6)$$

$$Var[Y_{ij}] = \sigma_s + \sigma_x + \sigma_r + \sigma_u. \quad (7)$$

$$\sigma_s = \lambda_{01}^2 * var[S_j] \text{ e } \sigma_x = b_{1j}^2 * var[X_{ij}]$$

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 σ_r e σ_u . O procedimento adotado neste trabalho para calcular a parte da variância atribuída ao diferentes níveis foi o seguinte.:

- a) em primeiro lugar, estimamos o modelo para o nível 1 (aluno), incluindo o conjunto de variáveis explicativas deste nível (a matriz X);
- b) em seguida, extraímos os valores dos componentes da variância para o intercepto e para o nível 1.

Nesta estimação com apenas um nível, o valor do componente da variância para o intercepto corresponde à parte da variância atribuída ao nível 2 (escolas), ou seja, a soma $(\sigma_x + \sigma_u)$ em (7); o valor do componente da variância para o nível 1 corresponde à parte da variância não explicada atribuída a este nível, após a inclusão das variáveis explicativas, representada por σ_r em (7). A razão $(\sigma_x + \sigma_u) / (\sigma_x + \sigma_u + \sigma_r)$ fornece a % da variância da proficiência dos alunos que é atribuída ao fator escola.

A equação (4) não é um modelo linear típico, tornando a estimação pelos mínimos quadrados ordinários não apropriada. Os termos de erro aleatórios, v_{ij} , 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 HLM⁴.

⁴ Mais detalhes sobre modelos hierárquicos podem ser encontrados em Bryk e Raudenbush (1992) e Raudenbush et alii, 2000.

4 - Fontes de dados e tratamento das variáveis

As três fontes de dados abrangidas nesta estimação são: Sistema Mineiro de Avaliação da Educação Pública (SIMAVE), Sistema Informacional Custo Aluno (SICA) e o Censo Escolar - 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 4ª e 8ª série do ensino fundamental e 3º ano do ensino médio das escolas públicas mineiras, bem como questionários destinados à comunidade escolar. As disciplinas avaliadas pelo PROEB são: Língua Portuguesa, Matemática, Ciências Humanas e Ciências da Natureza. Essas disciplinas são analisadas alternadamente. Dessa forma, em 2000 avaliaram-se Língua Portuguesa e Matemática; em 2001, Ciências da Natureza e Ciências Humanas; em 2002, Língua Portuguesa; e em 2003, Matemática.

As variáveis do SIMAVE são utilizadas no primeiro nível de análise do modelo hierárquico, o aluno, que, além disso, leva em consideração os efeitos espaciais por meio de *dummies* regionais. São 12 onze 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.

Do SIMAVE, portanto, extraímos:

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

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

Regmatic: 1 – aluno regularmente matriculado, 0 - caso contrário;

Esc_resp_fem: 1 - responsável feminina tem 8ª série completa ou mais, 0 - caso 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

Tarde: 1 - estuda no horário da tarde, 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 domicílios (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, essas 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 ensino⁵ Educação Infantil, Ensino Fundamental e Ensino Médio, nas modalidades Regular, Especial e Suplência, 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.

As demais variáveis da escola são retiradas do Censo Escolar. Esta base de dados é realizada e 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;

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

Proprofsup: 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.

Infra-estrutura: Índice que representa a escassez de infra-estrutura, calculado por meio da aplicação de uma Análise em Componentes Principais (ACP) às variáveis categóricas, ausência de abastecimento geral de água, de esgotamento sanitário e de coleta pública de lixo. A ACP é normalmente preferida pelos pesquisadores para compor índices a partir de variáveis que não podem ser exatamente mensuradas, pelo fato de fornecer uma solução matemática exata (Dunteman, 1989)⁶.

⁵ O banco compreende as escolas da área urbana e rural do estado de Minas Gerais.

⁶ Cabe aqui uma observação a respeito do uso da ACP para compor índices a partir de variáveis categóricas. A ACP é realizada a partir de uma matriz de covariâncias ou correlações de variáveis assumidas contínuas e normalmente distribuídas. Com variáveis categóricas ordenadas e/ou dicotômicas, a matriz de correlação de Pearson não é mais apropriada para o cálculo dos componentes. Quando as variáveis dicotômicas podem ser presumidas como medidas imperfeitas de variáveis subjacentes contínuas e latentes, alguns autores recomendam o uso de correlações *tetrachoric* (ou *polichoric*) na ACP (Harris, 1988).

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 segundo nível de análise, incorporamos, ainda, a proficiência média da escola (retirada do SIMAVE) como variável independente. A proficiência média da escola pode funcionar como uma *proxy* para várias características fixas e não mensuradas da escola como por exemplo, a excelência pedagógica, o currículo utilizado, aspectos administrativos, e mesmo outros fatores associados ao contexto social no qual a escola está inserida. Esses fatores fixos associados à escola podem também servir como um controle para o desempenho anterior do aluno, ou seja, o fato de um aluno frequentar determinada escola teria um efeito específico sobre o seu desempenho (Thomas e Mortimore, 1996).

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 o questionário. Entretanto, o número de observações existentes é suficiente para estimar os determinantes do desempenho educacional. A tabela seguinte mostra a redução total de observações por etapa educacional nos dois níveis de análise:

Tabela 1: Número de observações nas bases de dados segundo nível de análise e etapa educacional.

Nível de análise	Etapa educacional	Nº total de observações	Nº de observações sem missing
Primeiro nível: indivíduo	4ª série	140667	94614
	8ª série	195333	91109
	3º ano	117711	72941
Segundo nível: escola	4ª série	2469	1635
	8ª série	2552	1949
	3º ano	1453	1404

Fonte: Bases de dados do SIMAVE/03, SICA/03 e Censo Escolar/03.

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 (273,26), seguida pela 8ª série (241,89) e pelo 4º ano do ensino fundamental (189,71). Provavelmente, este resultado deriva, por um lado, do caráter abstrato desta disciplina, o que dificulta o aprendizado dos alunos mais jovens

e, por outro lado, da capacidade de ensino dos professores. Um professor de 4ª série ministra todo o conteúdo programático (Matemática, Português, Ciências, História e Geografia), não se constituindo, assim, em um especialista em Matemática, o mesmo não é esperado para professores da 8ª série e 3º ano que devem ser graduados na ciência.

Tabela 2: proficiência média dos alunos segundo nível de ensino - 2003.

Nível de ensino	Proficiência média	Desvio padrão
4ª série	189.71	48.44
8ª série	241.89	49.20
3º ano	273.26	50.29

Fonte: Elaboração própria a partir da base de dados do 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 e 3º ano (TAB. 3). À exceção do ensino médio, 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. Cerca de 60% dos alunos são regularmente matriculados na 4ª série e no 3º ano. Na última série do ensino fundamental, a proporção é bem mais baixa (39,13%).

A escolaridade da mãe é baixa, uma vez que mais de 60% não tem a 8ª série completa e, à medida que se avança na escolaridade do aluno, esta proporção aumenta. As responsáveis pelos alunos do 3º ano com escolaridade inferior a 8ª série completa compreendem 75,09% da amostra. Tal distribuição é esperada, posto que essas mães são mais velhas relativamente aos alunos da 4ª e 8ª séries e, portanto, sujeitas, a uma distribuição da educação mais concentrada nas primeiras séries do ensino fundamental.

Embora apenas um quarto dos domicílios de alunos tenha mais de 20 livros, acima de 70% dos alunos lê livros e isto é tanto mais evidenciado quanto menor o nível de escolaridade. No caso da leitura de jornais, ocorre o inverso. A incidência é maior, aproximadamente 66% dos alunos do ensino médio realizam este tipo de leitura.

Menos de um terço dos alunos nas três séries abordadas neste trabalho realizam 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.

Na 4ª série, mais de 70% fazem o dever de Matemática e essa proporção cai até o 3º ano, 40,73%. 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 proporção, 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		8ª série		3º ano	
	Sim	Não	Sim	Não	Sim	Não
Branco e amarelo	41.25	58.75	47.18	52.82	51.38	48.62
Regularmente matriculados	60.78	39.22	39.13	60.87	62.31	37.69
Responsável escolar feminina tem 8ª série completa	39.66	60.34	23.18	76.82	24.91	75.09
Possui mais de vinte livros em casa	25.51	74.49	27.44	72.56	29.78	70.22
Lêem livros	80.10	19.90	70.72	29.28	70.38	29.62
Lêem jornais	47.21	52.79	47.32	52.68	65.72	34.28
Não trabalham fora	85.64	14.36	66.25	33.75	49.40	50.60
Não trabalham em casa	29.93	70.07	27.66	72.34	27.47	72.53
Fazem dever de matemática	73.06	26.94	54.74	45.26	40.73	59.27
Nunca foram reprovados	77.44	22.56	69.03	30.97	64.12	35.88
Estudam de manhã	59.23	40.77	62.33	37.67	40.51	59.49
Estudam à tarde	40.34	59.66	22.73	77.27	2.57	97.43
Estudam à noite	0.37	99.63	14.89	85.11	56.87	43.13

Fonte: Elaboração própria a partir da base de dados do 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. A presença de laboratório é menor nas escolas que ofertam a 4ª série (31,73%) e maior nas de 3º ano do ensino médio (45,91%).

Tabela 4: Variáveis da escola: proporções segundo nível de ensino - 2003.

Variáveis	4ª série		8ª série		3º ano	
	Sim	Não	Sim	Não	Sim	Não
Recebem FNDE	80.47	19.53	78.88	21.12	75.91	24.09
Têm laboratório	31.73	68.27	39.46	60.54	45.91	54.09

Fonte: Elaboração própria a partir da base de dados do Censo Escolar/2003.

A proporção de professores com ensino superior é maior nas escolas de 8ª série (64%) e bem menor nas de 4ª série (34%), de acordo com a TAB. 5. O valor do custo dentro médio mensal por aluno é praticamente o mesmo para as séries do ensino fundamental (R\$ 62,00) e inferior no 3º ano do ensino médio (R\$ 40,50), devido às externalidades geradas nesta modalidade, pois as turmas são maiores (34,74 contra 28,88 alunos na 4ª série) e os professores tendem a lecionar em muitas turmas.

A variável ausência de infra-estrutura (devendo ser analisada em módulo) mostra que as escolas que ofertam o 3º ano estão mais bem providas de infra-estrutura urbana vis-à-vis as demais. Isto reflete o fato de a maioria delas estar concentrada nesta área geográfica.

Tabela 5: Variáveis da escola: valores médios segundo nível de ensino - 2003.

	4ª série	8ª série	3º ano
Proporção de professores ensino superior	0.34 (0.31)	0.64 (0.38)	0.51 (0.44)
Custo dentro da escola*	62.99 (98.29)	62.65 (105.13)	40.50 (17.65)
Ausência de infra-estrutura	-0.69 (0.76)	-0.85 (0.63)	-0.95 (0.42)
Matrículas por turma*	28.88 (6.88)	33.19 (7.31)	34.74 (7.40)

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

-Desvios padrões entre parênteses.

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

Por fim, há uma concentração de escolas nas três séries na Região Metropolitana de Belo Horizonte, Norte de Minas e Sul/Sudoeste e menor oferta na Central Mineira, Noroeste de Minas e Campo das Vertentes.

Tabela 6: Proporção de alunos por mesorregião segundo níveis de ensino - 2003.

Mesorregião	4ª série	8ª série	3º ano
Noroeste de Minas	2.49	2.02	2.32
Norte de Minas	12.77	11.80	9.22
Jequitinhonha	7.05	5.48	3.48
Vala do Mucuri	3.35	2.65	2.00
Triângulo Mineiro/Alto Paranaíba	9.14	10.13	9.28
Central Mineira	1.91	2.44	2.39
Metropolitana de Belo Horizonte	30.10	25.24	32.70
Vale do Rio Doce	9.03	8.94	8.03
Oeste de Minas	3.62	5.36	5.44
Sul/Sudoeste de Minas	8.79	12.82	12.21
Campo das Vertentes	2.23	2.90	3.27
Zona da Mata	9.52	10.22	9.66

Fonte: Elaboração própria a partir da base de dados do SIMAVE/2003.

5. 2 - Resultados dos modelos hierárquicos.

Os modelos hierárquicos estimados são especificados da seguinte forma:

Nível 1 (aluno):

$$Ln\ Profic = B0 + B1*(COR) + B2*(ESC_RESP) + B3*(LIVROS) + B4*(LER_LIVR) + B5*(LER_JORN) + B6*(NÃO TRAB_DOM) + B7*(NÃO TRAB_FOR) + B8*(DEVERMAT) + B9*(NREPROVA) + B10*(NORO) + B11*(NORTE) + B12*(MUCURI) + B13*(TRIANG) + B14*(CENTRAL) + B15*(METROP) + B16*(RDOCE) + B17*(OESTE) + B18*(SUL) + B19*(CVERT) + B20*(ZMATA) + B21*(MANHA) + B22*(NOITE) + B23*(REGMATRI) + R$$

Nível 2 (escola):

$$B0 = G00 + G01*(PROFIC3) + G02*(FNDE) + G03*(CDENTRO) + G04*(LABORATO) + G05*(MATTURMA) + G06*(MATTURM1) + G07*(PROPPROF) + G08*(INFRAEST) + U0$$

Na estimação do desempenho do aluno no 3º ano do ensino médio, não foram incluídas as variáveis Proporção de Professores no Ensino Superior (PROPPROF) devido ao elevado número de observações *missing* e a variável Infra-estrutura (INFRAEST) em decorrência de alta correlação com outras variáveis.

Os resultados dos modelos hierárquicos para a 4ª e 8ª séries do fundamental e 3º ano do ensino médio estão apresentados nas tabelas 7, 8 e 9, respectivamente. Como se observa, a maioria das variáveis é significativa, mas a importância delas altera conforme a série escolar. Ademais, em todos os modelos, as características ligadas aos alunos explicam uma parcela muito maior da variabilidade observada na proficiência em matemática do que as características das escolas.

No modelo para a quarta série do fundamental (TAB.7), os atributos da escola são responsáveis por 18,7% da variabilidade na proficiência em matemática, e as características dos alunos respondem pelos 80% restantes.

Quando se incluem as variáveis explicativas do nível 2, isto é, os atributos das escolas, a parcela não explicada da variância cai para 7,65%. Provavelmente, variáveis escolares não incluídas no modelo são importantes para explicar a proficiência em matemática dos alunos. Certamente, a localização de escolas em áreas urbanas, modernização dos equipamentos e presença de biblioteca devem contribuir para um melhor desempenho em Matemática. No entanto, estas variáveis não foram consideradas no modelo devido à elevada correlação com o índice construído para infra-estrutura.

Entre as variáveis de nível 2 incluídas no modelo, a proficiência média da escola é de longe a mais significativa. Entretanto, sua importância decresce à medida que se avança nos anos, isso porque, na 4ª série, um aumento de um desvio-padrão de 48,44 (TAB. 5) na proficiência média da escola está relacionado a um aumento de 4,2%⁷ na proficiência do aluno (TAB. 7), na 8ª série, um aumento de um desvio-padrão de 49,20 (TAB. 5) está relacionada a um aumento de 3,5% (TAB. 8) e, no 3º ano, um aumento de um desvio-padrão de 50,29 (TAB. 5) a 2,9% (TAB.9).

A outra variável significativa do nível da escola para a 4ª série é o tamanho das turmas (Mat_turma). Apesar do pouco efeito sobre o desempenho, seu sinal positivo mostra que, quanto maior a turma, melhor o desempenho. Tal evidência ratifica os achados de Hanushek (1998) que turmas pequenas não geram efeito nos anos escolares seguintes ao Jardim de Infância e de Krueger (1999) que a maior vantagem no atendimento em classes menores acontece no primeiro ano escolar, sendo o impacto menor nas séries seguintes, mas continua significativo.

Ainda sobre a estimação para a 4ª série (TAB. 7), à exceção das *dummies* de região e turno da noite, 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. A cor (branca ou amarela) aumenta em 14,5% a proficiência em matemática;
2. Nunca ter sido reprovado amplia em cerca de 10% o rendimento em matemática;
3. Ler livros melhora em 7,2 % o desempenho em matemática;
4. A responsável feminina com 8ª série completa ou mais aumenta em 4,6% a proficiência em matemática;
5. Fazer o dever de matemática melhora em, aproximadamente, 4,5% o rendimento na disciplina;
6. Não executar alguma tarefa doméstica ou trabalhar fora de casa aumenta respectivamente em 3,7% e 4%;
7. Quanto às mesorregiões do estado de Minas Gerais, quase todas tem uma proficiência menor em relação à região de referência (Jequitinhonha). A única positiva significativa é a Região Metropolitana (Metropolitana) cujos alunos apresentam um rendimento um pouco melhor, cerca de 2% comparativamente ao Jequitinhonha.

⁷ Para calcularmos o efeito sobre a variável dependente, utilizamos o antilog dado por $[\exp(\beta)-1] \times 100$.

Tabela 7: Resultados do Modelo Hierárquico para a 4ª série do ensino fundamental

Variável Dependente:	LNPROFIC	
Variáveis Independentes:	Coefficiente	T de student
<u>Nível II (nº observações: 1627)</u>		
Intercepto	4.9161	406.93
Fnde	-0.0016	-0.32
Custo_dentro	-0.00002	-1.65
Laboratório	0.0005	0.12
Prop_profsup	0.0033	0.38
Mat_turma	0.0010	3.88
Infraestrutura	-0.0018	-0.55
Proficiência_média escola	0.0041	44.04
<u>Nível I (nº observações: 94583)</u>		
Cor	0.1350	38.87
Esc_resp_fem	0.0450	9.35
Reg_matric	0.0372	6.95
Livros	0.0081	1.55
Ler_livros	0.0697	12.45
Ler_jornais	0.0343	7.61
Não Trab_domestico	0.0365	7.62
Não Trab_fora	0.0388	6.19
Dever_mat	0.0439	8.39
N_reprovado	0.0960	17.28
Turno_manhã	0.0209	2.92
Turno_noite	0.1069	1.00
Regiões		
Noroeste	-0.0187	-2.09
Norte	-0.0082	-1.19
Mucuri	0.0097	1.17
Triângulo	-0.0160	-2.04
Central	-0.0074	-0.67
Metropolitana	0.0190	2.72
Rio Doce	-0.0061	-1.01
Oeste	-0.0654	-2.92
Sul	-0.0185	-2.39
Campo das Vertentes	-0.0225	-2.12
Zona da Mata	-0.0126	-1.70
Componentes da Variância		
<u>a. Modelo com os dois níveis</u>		
Intercepto (escolas)	0.00407	
Nível I (alunos)	0.04914	
<u>b. Modelo com o nível I</u>		
Intercepto (escolas)	0.01133	
Nível I (alunos)	0.04916	

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

No modelo para a 8ª série, apresentado pela Tabela 8, a importância dos fatores escolares (9,86%) cai para quase a metade do observado para a 4ª série.

A variável de custo (Custo_dentro) é significativa apenas para a 8ª série, porém seu efeito sobre o desempenho é inexpressivo. Estimamos os modelos apenas com esta variável, buscando identificar a relação entre disponibilidade de recursos financeiros e proficiência do aluno e a variável

“Custo_dentro” só é significativa a 15%. Ao incluímos, também, proporção de professores com nível superior, a variável de custo passa a ser significativa a 8%.⁸

Cabe destacar que, nos resultados da estimação final, a proporção de professores com nível superior apresenta efeito negativo sobre a proficiência do aluno da 8ª série (TAB. 8). Este resultado não é pertinente com as evidências dos estudos sintetizados na seção dois que, ora mostram a importância da capacitação como os de Riani (2005) e de Albernaz *et alii* (2002), ora mostram menor relevância (BARROS *et alli*, 2001; HANUSHEK *et alli*, 1996), mas não um efeito contrário.

As variáveis do primeiro nível, referentes ao aluno, são todas significativas e positivas, com exceção da variável turno noturno, evidenciando uma queda de cerca de 1%. Por ordem de importância em termos dos determinantes do desempenho em Matemática, verificamos que:

1. O fato de nunca ter sido reprovado (N_reprovado) é a variável de maior peso na explicação da 8ª série, pois responde por aumento de 8,4%;
2. Fazer o dever de matemática contribui com melhora de 4% no desempenho do aluno;
3. Tanto estar regularmente matriculado como a mãe possuir 8ª série completa ou mais aumentam em cerca de 3% a proficiência;
4. O domicílio contar com mais de 20 livros (Livros) amplia a proficiência em 2,7%;
5. A não realização de tarefas domésticas (Não Trab_doméstico) aumenta o desempenho em 2,4%;
6. Ler livros melhora em 1,9%;
7. A cor branca ou amarela tem um efeito de 1,5% sobre o rendimento escolar.

No caso das *dummies* regionais, o desempenho médio dos alunos residentes nas mesorregiões é inferior ao da categoria de referência, a região Vale do Jequitinhonha. Nas regiões mais pobres, existem poucas escolas da rede privada. Por isso, os estabelecimentos da rede estadual recebem alunos provenientes de famílias de vários estratos de renda, o que contribui para melhorar o desempenho médio da escola e da região. Ademais, é a meso de menor desvio-padrão da proficiência em Matemática nas três séries avaliadas (Tabelas em Anexos).

⁸ Em razão destes resultados pouco elucidativos, optamos por não apresentá-los no texto.

Tabela 8: Resultados do Modelo Hierárquico para a 8ª série do ensino fundamental

Variável Dependente:	LNPROFIC	
Variáveis Independentes:	Coefficiente	t de student
Intercepto	5.3261	1429.15
Fnde	0.0013	0.96
Custo_dentro	-0.00003	-7.56
Laboratório	0.0012	1.09
Prop_profsup	-0.0119	-3.53
Mat_turma	-0.0002	-2.80
Infraestrutura	-0.0001	-0.08
Proficiência_média escola	0.0034	94.54
<u>Nível I (nº observações: 91078)</u>		
Cor	0.0153	11.93
Esc_resp_fem	0.0337	20.97
Reg_matric	0.0308	21.16
Livros	0.0268	18.50
Ler_livros	0.0190	13.02
Ler_jornais	0.0124	9.39
Não Trab_domestico	0.0238	16.06
Não Trab_fora	0.0094	6.97
Dever_mat	0.0354	26.65
N_reprovado	0.0804	46.22
Turno_manhã	0.0094	4.92
Turno noite	-0.0100	-3.69
<u>Regiões</u>		
Noroeste	-0.0154	-3.98
Norte	-0.0146	-5.36
Mucuri	-0.0118	-2.94
Triângulo	-0.0182	-5.87
Central	-0.0067	-1.88
Metropolitana	-0.0067	-2.31
Rio Doce	-0.0076	-2.76
Oeste	-0.0090	-2.72
Sul	-0.0082	-2.79
Campo das Vertentes	-0.0102	-2.67
Zona da Mata	-0.0046	-1.62
Componentes da Variância		
<u>a. Modelo com os dois níveis</u>		
Intercepto (escolas)	0.00000	
Nível I (alunos)	0.03144	
<u>b. Modelo com o nível I</u>		
Intercepto (escolas)	0.00347	
Nível I (alunos)	0.03170	

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

No 3º ano do ensino médio, a importância dos fatores escolares é de 10,65% na explicação do desempenho do aluno em Matemática (TAB. 9). Embora matrícula por turma e proficiência média da escola sejam significativas, o efeito é praticamente nulo.

Entretanto, as variáveis do 1º nível são quase todas significativas e positivas, à exceção e do turno da noite que possui sinal negativo, representando uma queda de 1,6% na proficiência (TAB. 9). As três variáveis de nível 1 mais importantes para explicar a proficiência em matemática no 3º ano são estar regularmente matriculado que acresce a proficiência em 5,2%; não realizar trabalho

doméstico, contribuindo com 4,82% e nunca ter sido reprovado que amplia em 4,7% o rendimento em Matemática.

Tabela 9: Resultados do Modelo Hierárquico para a 3º ano do ensino médio

Variável Dependente:	LNPROFIC	
Variáveis Independentes:	Coefficiente	t de student
<u>Nível II (nº observações: 1398)</u>		
Intercepto	5.4967	890.51
Fnde	0.00001	0.01
Custo_dentro	0.00002	0.36
Laboratório	0.0014	1.13
Mat_turma	-0.0005	-5.18
Proficiência média escola	0.0029	68.79
<u>Nível I (nº observações: 72912)</u>		
Cor	0.0080	6.14
Esc_resp_fem	0.0194	12.77
Reg_matric	0.0506	29.14
Livros	0.0176	12.19
Ler_livros	0.0006	0.43
Ler_jornais	0.0186	14.79
Não Trab_domestico	0.0471	32.72
Não Trab_fora	0.0049	3.74
Dever_mat	0.0251	19.64
N_reprovado	0.0457	27.07
Turno_manhã	0.0133	2.48
Turno_noite	-0.0162	-3.10
<u>Regiões</u>		
Noroeste	-0.0249	-6.17
Norte	-0.0157	-4.29
Mucuri	-0.0120	-2.44
Triângulo	-0.0237	-6.33
Central	-0.0181	-3.75
Metropolitana	-0.0149	-4.22
Rio Doce	-0.0108	-3.01
Oeste	-0.0163	-4.20
Sul	-0.0168	-4.49
Campo das Vertentes	-0.0087	-1.79
Zona da Mata	-0.0113	-2.96
Componentes da Variância		
<u>a. Modelo com os dois níveis</u>		
Intercepto (escolas)	0.00001	
Nível I (alunos)	0.02518	
<u>b. Modelo com o nível I</u>		
Intercepto (escolas)	0.00289	
Nível I (alunos)	0.02531	

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

A exemplo do modelo da 8ª série, o desempenho médio dos alunos residentes em todas as “meso” é inferior ao da região do Jequitinhonha. Neste caso, os argumentos apresentados para a análise da 8ª série são ainda mais robustos, por se tratar da última série do Ensino Médio. É praticamente restrita à rede pública estadual a oferta de vagas no nível médio.

6 – Considerações finais

Os resultados deste trabalho constituem-se em mais uma evidência de que o efeito da escola 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 permitirem a separação de efeitos de componentes de vários níveis e, por isso, tendem a não subestimar o efeito da escola sobre o desempenho do aluno.

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 baixo. Incluímos, também, a proficiência média da escola como variável independente para contornar a ausência de várias características fixas e não mensuradas da escola, já citadas, como a excelência pedagógica, o currículo utilizado, aspectos administrativos, e mesmo outros fatores associados ao contexto social no qual a escola está inserida que, embora tenha sido a de maior poder de explicação, não foi suficiente para minorar o efeito das variáveis referentes ao aluno.

Deste modo, em termos das características do aluno, chama atenção, em todas as séries analisadas, o efeito positivo da não-reprovação, de estar regularmente matriculado e do não-trabalho em casa sobre o desempenho em Matemática. No caso das séries do ensino fundamental, além desses efeitos, são importantes e comuns os impactos positivos da escolaridade da mãe, da execução do dever de Matemática, da leitura de livros e do não trabalho dentro ou fora de casa. Por outro lado, o efeito mais importante para a 4ª série é a cor, sobressaindo os brancos e amarelos que perfazem apenas 40% dos alunos. Para a 8ª série, um efeito específico, apesar de não ser o mais importante, é a presença de uma mini-biblioteca em casa.

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 treinamento de professores e na modernização de equipamentos, certamente, promoverá maior aproveitamento do aluno.

Referências Bibliográficas

- AFONSO, A. E AUBYN, M.. “Cross-Country Efficiency of Secondary Education Provision: A semi-parametric Analysis with non-discretionary Inputs” **European Central Bank Working Paper Series** No. 494/June 2005. <http://www.ecb.int/pub/pdf/scpwps/ecbwp494.pdf> acesso 27/02/2006.
- ALBERNAZ, A., FERREIRA, F. H. G., FRANCO, C. Qualidade e Equidade na Educação Fundamental Brasileira. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.32, n.3, p.453-476, dez. 2002.
- BARBOSA, M. E. F, FERNANDES, C. A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4ª série. In: FRANCO, C. (Org.) **Avaliação, ciclos e promoção na educação**. Porto Alegre: ArtMed, . p.155-172, 2001.
- BARROS, R. P. *et al.* Determinantes do desempenho educacional do Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.31, n.1. p.1-42, abr. 2001.
- CÉSAR, C. C., SOARES, J. F. Desigualdades acadêmicas induzidas pelo contexto escolar. **Revista Brasileira de Estudos Populacionais**, v.18, n.1/2, jan./dez, p.97-110, 2001.
- DUNTEMAN, G. H. *Principal components analysis*. Thousand Oaks, CA: Sage Publications, Quantitative Applications in the Social Sciences Series, No. 69, 1989.
- FELÍCIO, Fabiana de.; FERNANDES, Reynaldo. *O Efeito da Escola sobre o Desempenho Escolar: Uma Avaliação do Ensino Fundamental no Estado de São Paulo*. Anais do XXXIII Encontro Nacional de Economia 156, ANPEC, 2005.
- FERRÃO, M. E. *et al.* O SAEB – Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. **Revista Brasileira de Estudos Populacionais**, v.18, n.1/2, p.111-130, jan./dez, 2001.
- FLETCHER, P. R. **À procura do ensino eficaz**. Rio de Janeiro, RJ: MEC, 1998.
- HANUSHEK, E. A, GOMES-NETO, J. B., HARBISON, R. W. Efficiency-enhancing investments in school quality. In: BIRDSALL, N., SABOT, R. H (Eds.) **Opportunity forgone: education in Brazil**. Washington, DC.: Inter-American development Bank. p.385-424,1996.
- HANUSHEK, E. A. **The evidence on class size**. Rochester, N.Y.: University of Rochester, W. Allen Wallis Institute of political Economy, 1998. (Occasional Paper, 98-1)
- HARRIS, B. *Tetrachoric correlation coefficient*. In Kotz L, Johnson NL (Eds.), *Encyclopedia of statistical sciences*. Vol. 9 (pp. 223-225). New York: Wiley, 1988.
- KRUEGER, A. B. Experimental estimates of education production functions. **The Quarterly Journal of Economics**, v.114, n.2, p.497-532, May. 1999
- RIANI, J. L. R. Determinantes do Resultado Educacional no Brasil: Família, Perfil Escolar dos Municípios e Dividendo Demográfico Numa Abordagem Hierárquica e Espacial. Belo Horizonte. UFMG/Cedeplar, 2005.
- RIOS-NETO, E. L. G., CÉSAR, C. C. e RIANI, J.L.R. **Estratificação educacional e progressão escolar por série no Brasil**. Artigo apresentado no encontro Determinantes do Sucesso Educacional. IPEA, Maio 2002.
- SILVA, N. V., HASENBALG, C. **Recursos familiares e transições educacionais**. Versão Preliminar apresentada no Workshop de Demografia da Educação. Associação Brasileira de Estudos Populacionais, ABEP, Salvador, Bahia, Junho 2001.

THOMAS e MORTIMORE, P. Comparison of Value-Added Models for Secondary-School Effectiveness. **Research Papers in Education**, vol. 11, n. 1, p. 5-23, 1996.

Anexos

Tabela A1: Proficiência por mesorregião - 4ª série

Mesorregião	Obs	Média	Desvio padrão	Min	Max
Noroeste de Minas	2295	200.36	53.48	86.70	342.30
Norte de Minas	13312	192.05	54.99	75.70	342.30
Jequitinhonha	7400	188.26	50.73	82.27	342.30
Vale do Mucuri	3475	185.37	49.60	80.82	338.18
Triângulo Mineiro/Alto Paranaíba	8472	216.64	55.60	75.48	338.18
Central Mineira	1836	209.96	53.88	89.51	338.18
Metropolitana de Belo Horizonte	28630	222.02	54.38	82.10	342.30
Vale do Rio Doce	9391	203.79	56.07	81.72	342.30
Oeste de Minas	3450	220.80	57.80	78.50	338.18
Sul/Sudoeste de Minas	6355	213.72	53.81	83.87	342.30
Campo das Vertentes	1931	220.77	54.00	84.39	342.30
Zona da Mata	8067	207.36	54.30	81.77	342.30

Fonte: Elaboração própria a partir da base de dados do SIMAVE/2003.

Tabela A2: Proficiência por mesorregião - 8ª série.

Mesorregião	Obs	Média	Desvio padrão	Min	Max
Noroeste de Minas	2196	241.33	46.33	115.70	401.82
Norte de Minas	11147	226.99	45.29	106.86	399.71
Jequitinhonha	5343	237.42	45.71	108.74	400.18
Vale do Mucuri	2287	231.24	46.62	104.38	403.55
Triângulo Mineiro/Alto Paranaíba	11215	252.43	47.51	109.87	403.55
Central Mineira	2666	249.78	47.86	113.60	403.55
Metropolitana de Belo Horizonte	14146	247.15	48.39	114.06	403.55
Vale do Rio Doce	9281	242.00	47.56	107.93	401.15
Oeste de Minas	5511	255.63	49.59	110.78	401.82
Sul/Sudoeste de Minas	14409	255.85	48.62	113.81	403.55
Campo das Vertentes	3062	252.20	49.82	113.01	401.82
Zona da Mata	9846	248.83	50.00	108.84	403.55

Fonte: Elaboração própria a partir da base de dados do SIMAVE/2003.

Tabela A3: Proficiência por mesorregião - 3º ano.

Mesorregião	Obs	Média	Desvio padrão	Min	Max
Noroeste de Minas	1962	263.31	45.93	156.25	450.13
Norte de Minas	7712	256.15	46.32	153.28	438.38
Jequitinhonha	3194	268.27	47.22	158.99	433.11
Vale do Mucuri	1463	266.10	48.06	159.93	438.38
Triângulo Mineiro/Alto Paranaíba	8076	284.93	50.04	157.29	450.13
Central Mineira	2082	280.75	48.64	156.95	444.49
Metropolitana de Belo Horizonte	19546	276.55	50.74	154.22	444.58
Vale do Rio Doce	6727	275.43	49.17	156.65	443.71
Oeste de Minas	4107	284.54	48.44	153.26	443.71
Sul/Sudoeste de Minas	9157	286.05	49.92	153.70	444.49
Campo das Vertentes	2816	287.38	52.14	161.41	450.13
Zona da Mata	6099	281.97	50.90	153.61	445.76

Fonte: Elaboração própria a partir da base de dados do SIMAVE/2003.