

FATORES RELACIONADOS AO ATRASO ESCOLAR NO ESTADO DE MINAS GERAIS*

JOSIANE SOUZA DE PAULA^I
ANA MARIA DE PAIVA FRANCO^{II}
JOSÉ WALDEMAR DA SILVA^{III}

RESUMO

O objetivo deste trabalho é analisar os fatores que se relacionam com o atraso escolar em Minas Gerais. A metodologia utilizada envolve a aplicação de modelos logísticos hierárquicos para os dados da Prova Brasil de 2013, referentes aos 5º e 9º anos, na tentativa de identificar se existem diferenças entre os fatores associados ao atraso escolar nos anos iniciais e finais do ensino fundamental. Os resultados obtidos mostram que características dos estudantes e também as escolares estão relacionadas com a probabilidade de ocorrência de atraso escolar, havendo diferenças regionais nesse aspecto. Em particular, observa-se que as maiores chances de defasagem idade-série são para alunos do sexo masculino, não brancos, que trabalham fora de casa, estudam em turmas heterogêneas e em escolas que adotam programas de redução de taxas de abandono. Além disso, fica caracterizado que o desempenho dos estudantes considerados em atraso escolar é mais prejudicado em algumas escolas do que em outras, dependendo do nível socioeconômico do aluno.

I Instituto de Economia e Relações Internacionais, Universidade Federal de Uberlândia (IERI/UFU), Uberlândia-MG, Brasil; josiane_udi@yahoo.com.br.

II Instituto de Economia e Relações Internacionais, Universidade Federal de Uberlândia (IERI/UFU), Uberlândia-MG, Brasil; ana.paiva@ufu.br.

III Faculdade de Matemática, Universidade Federal de Uberlândia (Famat/UFU), Uberlândia-MG, Brasil; zewaldemar@ufu.br.

* A autora Josiane Souza de Paula agradece à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) o apoio financeiro.

PALAVRAS-CHAVE REPETÊNCIA • MODELO LOGÍSTICO HIERÁRQUICO • PROVA BRASIL • MINAS GERAIS.

FACTORES RELACIONADOS CON EL ATRASO ESCOLAR EN EL ESTADO DE MINAS GERAIS

RESUMEN

El objetivo de este trabajo es analizar los factores que se relacionan con el atraso escolar en Minas Gerais. La metodología utilizada abarca la aplicación de modelos logísticos jerárquicos para los datos de la Prova Brasil del 2013, relativos a los 5º y 9º años, con el intento de identificar si hay diferencias entre los factores asociados al atraso escolar en los primeros y los últimos años de la educación fundamental. Los resultados que se obtuvieron muestran que las características de los estudiantes y también las escolares se relacionan con la probabilidad de que ocurra un atraso escolar, habiendo diferencias regionales en este aspecto. Por encima de todo se observa que las mayores chances de desfase edad-año escolar ocurren con alumnos del sexo masculino, no blancos, que trabajan fuera de casa, estudian en grupos heterogéneos y en escuelas que adoptan programas de reducción de tasas de abandono. Además, se caracteriza que el desempeño de los estudiantes considerados en atraso escolar es más perjudicado en algunas escuelas que en otras en función del nivel socioeconómico del alumno.

PALABRAS CLAVE REPETICIÓN • MODELO LOGÍSTICO JERÁRQUICO • PROVA BRASIL • MINAS GERAIS.

FACTORS RELATED TO SCHOOL DELAY IN THE STATE OF MINAS GERAIS

ABSTRACT

This article aims to analyze the main factors related to school delay in the State of Minas Gerais. We applied Hierarchical Logistic Models, based on data from Prova Brasil 2013 for 5th and 9th graders, to identify whether there are differences between the factors associated with school delay in the early and final years of elementary education. Results show that student characteristics and school factors are related to the probability of school retention, with regional differences in this regard. Results also show that the greatest chance of age-grade mismatch is for non-white male students, who also work part- or full-time and attend heterogeneous classes in schools with programs to reduce dropout rates. In addition, the performance of retained students seems to be harmed more in some schools than in others, depending on students' socioeconomic level.

KEYWORDS GRADE REPETITION • HIERARCHICAL LOGISTIC MODEL • PROVA BRASIL • MINAS GERAIS.

INTRODUÇÃO

O desempenho escolar dos estudantes pode divergir devido a diversos fatores, entre os quais encontra-se o atraso escolar. Considerando uma trajetória regular, no Brasil o aluno deve iniciar o ensino fundamental aos seis anos de idade e completá-lo aos 14 anos, ingressando no ensino médio aos 15 anos e concluindo aos 17 anos. O aluno que apresenta dois anos ou mais de diferença da idade recomendada para a série que cursa está em atraso escolar, contribuindo para a origem de imperfeições no fluxo escolar. A demora na matrícula, abandono ou reprovação estão entre as principais causas do atraso escolar. Não ter a idade esperada para a série causa desvantagens aos estudantes, como, por exemplo, afeta a qualidade do aprendizado e compromete a autoestima (HONDA, 2006). O atraso escolar pode ser determinado por características individuais, da família, do meio social e do ambiente escolar (PONTILI; KASSOUF, 2007). Tendo isso em mente, questiona-se quais fatores estão relacionados com a probabilidade de um estudante estar em atraso escolar e em que sentido as escolas são importantes para evitar que esse fato ocorra.

O governo brasileiro, considerando a relevância da escolaridade para a renda nacional e o bem-estar social, promoveu várias reformas, principalmente, nos anos 1990, tendo em vista a melhoria dos índices educacionais, entre os quais estão os indicadores de acesso e fluxo escolar (PONTILI; KASSOUF, 2007). Essas medidas representaram uma alteração de enfoque das políticas educacionais, com a vinculação de parte das receitas públicas aos gastos em educação, em particular no ensino fundamental, e com um nível mínimo de gasto por estudante, além de modificações na duração do ano letivo e na qualificação mínima exigida para os professores (VASCONCELLOS, 2005).

Nesse sentido, a educação brasileira passou por expressivas mudanças nas últimas décadas, principalmente em termos de melhorias de acesso ao ensino fundamental.¹ Em 2006, por exemplo, as crianças passaram a ingressar no ensino fundamental com duração de nove anos (o limite final para implementação da lei foi em 2010) aos seis anos de idade e, em 2009, a escolaridade obrigatória, que se limitava a esse nível de ensino, englobou também o ensino médio. Destaca-se que Minas Gerais antecipou a política de aumento da duração do ensino fundamental de oito para nove anos, sendo o primeiro estado a aceitar crianças de seis anos no ensino fundamental.

Mesmo diante dos avanços observados, o atraso escolar ainda é um problema que compromete a qualidade da educação e a superação das desigualdades existentes na sociedade brasileira, haja vista a relação entre qualidade da educação e renda futura (CURI; MENEZES-FILHO, 2009).

O atraso escolar difere conforme a realidade econômica e social dos estados e regiões (PONTILI; KASSOUF, 2007). De acordo com os dados do Censo Escolar, a taxa de distorção idade-série no Brasil reduziu-se de 28,6%, em 2006, para 19,2%, em 2015. Entre as regiões, a maior taxa, em 2015, correspondia ao Norte (28,3%) e a menor pertencia ao Sudeste (12,7%). A maior proporção de alunos em atraso escolar é encontrada nas áreas rurais, principalmente, na rede pública de ensino e nos anos finais do ensino fundamental. Em Minas Gerais, a taxa de distorção idade-série no ensino fundamental

¹ Além da questão normativa, outro fator apontado para a melhoria do acesso à escola foi o da "transição demográfica", haja vista a redução do crescimento populacional (HONDA, 2006), principalmente de crianças e jovens (RIANI, 2005).

foi de 12,6%, nos anos iniciais foi de 5,3% e nos anos finais abrangeu 20,6% dos estudantes (BRASIL, 2015).

Segundo dados do Instituto Brasileiro de Geografia e Estatística (IBGE) (2016), o atraso escolar afetou 40,7% dos alunos mais carentes no Brasil em 2015. Entre os 20% mais pobres, a taxa de distorção idade-série foi cinco vezes maior do que entre os alunos 20% mais ricos conforme a distribuição do rendimento mensal domiciliar *per capita*. Além disso, observou-se que 1,6 milhão de jovens entre 15 e 17 anos de idade não estudavam, dos quais 61,4% tinham o ensino fundamental incompleto e 16,4% abandonaram a escola sem concluir o ensino médio, ou seja, grande parte da evasão escolar ocorre no ensino fundamental, atingindo, principalmente, os jovens mais fragilizados economicamente.

Diante dessas informações, questiona-se sobre quais variáveis podem estar relacionadas à trajetória escolar dos estudantes. No presente artigo, o objetivo é identificar as principais características individuais, familiares e escolares que estão ligadas à probabilidade de atraso escolar nas escolas públicas do estado de Minas Gerais para o 5º e 9º anos do ensino fundamental. Acredita-se que muitos dos resultados gerais podem ser verificados em outras partes do Brasil. A variável dependente (ou resposta) do modelo refere-se à condição de atraso escolar do aluno, sendo essa uma variável binária, que assume o valor 1 se o estudante tem dois anos ou mais de diferença da idade adequada para a série que frequenta, e zero, caso contrário. Considerando a estrutura hierárquica existente nos dados educacionais, o modelo escolhido foi o *logit* hierárquico (ou multinível).

A preocupação com o atraso escolar deriva de questões orçamentárias e sociais. A importância de se analisarem os motivos vinculados ao elevado atraso escolar se justifica pela possibilidade de, por meio dessa análise, se vislumbrarem pontos de ação para políticas públicas que sejam capazes de reverter esse oneroso problema educacional. Mesmo que tal tema já venha sendo discutido (HONDA, 2006; PONTILLI; KASSOUF, 2007), a possibilidade de utilizar novas bases de dados e explorar técnicas estatísticas atuais justificam seu estudo.

O artigo está dividido em seis seções, incluindo esta introdução. A seguir é feita uma revisão da literatura sobre fatores associados à trajetória escolar dos jovens. Posteriormente apresentam-se dados da educação básica no estado de Minas Gerais e são abordados o modelo logístico hierárquico, as fontes de dados e as estatísticas descritivas. Discutem-se os resultados do modelo estimado e, por fim, são tecidas as considerações finais.

FATORES ASSOCIADOS À TRAJETÓRIA ESCOLAR

Considerando-se os benefícios econômicos e sociais da educação, identificar as principais origens dos problemas escolares tem sido agenda de pesquisa de gestores e estudiosos (LEON; MENEZES-FILHO, 2002; PONTILI; KASSOUF, 2007; SHIRASU; ARRAES, 2015; SOARES, 2002). A pesquisa empírica tem procurado estudar a trajetória e o desempenho escolar dos jovens, buscando, basicamente, compreender quais os fatores que contribuem para um adequado fluxo no sistema regular de ensino, bem como aqueles que são responsáveis pelo avanço do aprendizado e, portanto, da qualidade escolar (BIONDI; FELÍCIO, 2007).

Gomes-Neto e Hanushek (1994), ao estudarem a reprovação de alunos das áreas rurais do Nordeste brasileiro, encontraram, entre os fatores mais importantes a ela associados, o desempenho do aluno e as políticas governamentais na escola (políticas de acesso e de melhoria na qualidade escolar).

Silva e Hasenbalg (2002), a partir do modelo logístico, analisando as características ligadas à progressão escolar, observaram que fatores socioeconômicos considerados em diferentes estágios do processo de escolarização têm efeito máximo no meio do ciclo escolar básico.

Nesse mesmo sentido, Rios-Neto, Cezar e Riani (2002), analisando a progressão por série, por meio de um modelo logístico hierárquico com dados da Pesquisa Nacional por Amostra de Domicílios (Pnad), verificaram que a escolaridade da mãe e a educação média do professor têm um efeito positivo na probabilidade de progressão escolar na 1ª série. No entanto, a partir da análise dos componentes aleatórios

do modelo, nota-se que a educação média do professor gera um efeito negativo na promoção através da escolaridade da mãe. Conclui-se que a escolaridade do professor (considerada uma *proxy* para a política de qualidade educacional) seria um substituto para a escolaridade da mãe, sendo que professores mais qualificados promoveriam a melhora do desempenho do aluno e a redução na desigualdade educacional.

Leon e Menezes-Filho (2002) desenvolveram um modelo para estudar reprovação, avanço e evasão escolar no Brasil, para os ensinos fundamental e médio, com dados da Pesquisa Mensal de Emprego (PME), do IBGE, considerando o período de 1984 a 1997. Os principais resultados indicam perda de importância das variáveis familiares ao longo do tempo, principalmente nas séries mais avançadas.

Já Ferrão, Beltrão e Santos (2002) aplicaram modelos hierárquicos para os dados do Sistema de Avaliação da Educação Básica (Saeb) 1999, referentes à 4ª série do ensino fundamental, para a região Sudeste, com o objetivo de avaliar o efeito de políticas de não repetência. Os resultados sugerem que o desempenho acadêmico dos estudantes com idade superior à adequada para a série frequentada é inferior ao dos alunos em idade adequada e, em algumas escolas, a proficiência é mais penalizada.

Riani (2005) avaliou a probabilidade de frequentar a escola na idade adequada e a probabilidade de progressão por série, utilizando modelos hierárquicos e espaciais para os ensinos fundamental e médio no Brasil. As análises da autora revelam que a educação da mãe é o fator associado mais relevante e que a trajetória escolar do estudante está fortemente relacionada à sua origem social, evidenciando a estratificação educacional existente no Brasil.

Com dados da Pnad de 1996, Honda (2006) observou que a educação dos pais é importante para reduzir o atraso escolar e que a renda familiar também exerce influência nesse aspecto. Por sua vez, Pontili e Kassouf (2007) analisaram o ensino fundamental, para os estados de São Paulo e Pernambuco. Os resultados sugerem que políticas voltadas para melhorar a escolaridade dos chefes de família e/ou a renda familiar podem reduzir o atraso e elevar a frequência

escolar. Já nas ações destinadas a melhorar a infraestrutura das escolas é preciso considerar as especificidades locais e a meta a ser atingida.

Do mesmo modo, Shirasu e Arraes (2015), utilizando um modelo logístico multinível para as escolas públicas do Ceará no período 2009 a 2011, analisaram a evasão e a repetência escolar no ensino médio. Os resultados indicam que as maiores chances de abandono ocorrem em virtude da repetência persistente, do desinteresse pelos estudos e do elevado atraso escolar.

Por fim, percebe-se que ainda há controvérsias nesse debate, principalmente pelo uso de diferentes bases de dados e pelo emprego de metodologias diversas, além da forma de tratamento das variáveis. Assim, espera-se com o presente estudo contribuir para as discussões acerca de um importante desafio para a educação brasileira, que é a elevada taxa de atraso escolar.

ANÁLISE DE DADOS EDUCACIONAIS EM MINAS GERAIS

No Brasil, com base nos dados da Pnad, a proporção dos estudantes com atraso no ensino fundamental foi mais elevada entre os estudantes da rede pública (28,7% contra 7,9% da rede particular), homens (31,6%), residentes em área rural (40%) e negros (31,4%) (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE, 2016).

Analisando a trajetória dos estudantes do 5º ano do ensino fundamental, em Minas Gerais (Tabela 1), observa-se que, na rede estadual, em 2013, 99% dos estudantes matriculados foram aprovados, 0,4% foram reprovados e 0,6% deixaram de frequentar a escola durante o ano letivo. Já para a rede municipal, os resultados foram inferiores, cujos índices corresponderam a 95% de aprovação, 3,6% de reprovação e 0,5% de abandono da escola. Em 2014, os resultados foram bastante similares, com 99% e 96,1% dos estudantes aprovados, respectivamente, nas redes estadual e municipal. Esses valores podem ser justificados pela adoção de sistemas de progressão continuada nos anos iniciais do ensino fundamental em Minas Gerais.

TABELA 1 - Trajetória dos estudantes do 5º ano do ensino fundamental, Minas Gerais

TAXA	2013		2014	
	ESTADUAL	MUNICIPAL	ESTADUAL	MUNICIPAL
Aprovação	99,0	95,9	99,0	96,1
Reprovação	0,4	3,6	0,5	3,4
Abandono	0,6	0,5	0,5	0,5
Distorção idade-série	9,5	18,9	7,0	15,6

² Disponível em: <<http://inepdata.inep.gov.br/>>. Acesso em: 11 ago. 2016.

Fonte: InepData, 2016.²

³ Por trás desse conceito estão diferentes aspectos da educação, como a idade de ingresso na escola, taxa de reprovação e abandono.

A taxa de distorção idade-série (uma medida de atraso escolar) é obtida por meio do cálculo da porcentagem de estudantes que estão com dois anos ou mais acima da idade recomendada para a série. Em Minas Gerais, essa taxa é de 7% para os alunos do 5º ano da rede estadual e 15,6% para os da rede municipal, em 2014.³ Assim, mesmo diante das melhorias observadas, as taxas de conclusão do ensino fundamental na idade adequada não são tão favoráveis. Conforme o Instituto de Pesquisa Econômica Aplicada (IPEA, 2014), menos da metade dos jovens com 15 anos havia concluído o ensino fundamental em 2013 e 20% das pessoas com 18 anos ainda não tinham o diploma desse nível de ensino.

A taxa de frequência escolar bruta no país, em 2015, atingiu 98,6% entre as crianças e adolescentes de 6 a 14 anos, um reflexo da intensificação da expansão da educação básica. No entanto, quanto às pessoas de 15 a 17 anos, essa taxa ficou em 85%, progredindo lentamente (crescimento de apenas 3,4 pontos percentuais em relação a 2005) devido à dificuldade de garantir educação básica obrigatória para essa faixa etária. A despeito das políticas de correção de fluxo praticadas, o país ainda apresenta altas taxas de repetência quando comparado aos países da América Latina com níveis elevados para os jovens de 15 anos, que são referência para a avaliação do Programa Internacional de Avaliação de Alunos (Programme for International Student Assessment – Pisa). Esse resultado leva o país a um desempenho inferior em relação a outras nações (IBGE, 2016). No entanto, destaca-se que somente a frequência escolar não indica o progresso

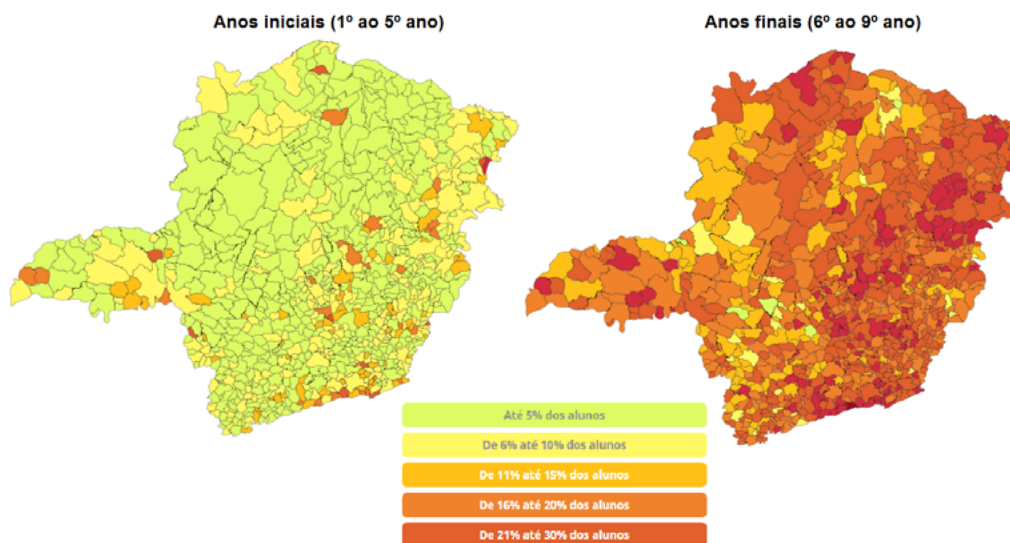
dos alunos na escola, pois são várias as razões que afetam negativamente o desenvolvimento do aluno e podem causar atraso durante o percurso escolar.

A Figura 1 mostra dados mais recentes da distorção idade-série para Minas Gerais. Observa-se que, em 2015, nos anos iniciais do ensino fundamental (1º ao 5º ano), grande parte dos alunos estava na idade adequada para a série, com uma taxa de distorção idade-série, incluindo escolas públicas, privadas, da área urbana e rural, em torno de 5%. No entanto, para os anos finais do ensino fundamental (6º ao 9º ano), essa taxa foi de 21%, sendo pouquíssimos os municípios (dentre os 853 existentes no estado) com taxa de até 5%, havendo casos em que mais da metade dos estudantes encontrava-se em atraso escolar.

Assim, quanto mais elevada a idade, maior é o custo de oportunidade de se estar na escola (VASCONCELLOS, 2005). Os dados apresentados revelam altas taxas de reprovação e abandono escolar de estudantes que ainda estão em idade em que a educação básica é obrigatória, principalmente para aqueles com menores condições econômicas, elevando-se, dessa forma, as desigualdades educacionais.

4 Disponível em: <<http://www.qedu.org.br/>>. Acesso em: 09 set. 2016.

FIGURA 1 - Distorção idade-série no ensino fundamental, Minas Gerais (2015)



Fonte: Inep (2015). Organizado por QEDU.⁴

O MODELO LOGÍSTICO HIERÁRQUICO

A variável resposta (Y) que representa o atraso escolar é uma *dummy*, a qual assume valores 0 ou 1. Nesse caso, a abordagem linear não é adequada porque o termo de erro do nível 1 não segue os pressupostos de normalidade e homogeneidade, uma vez que o valor predito é considerado em termos de probabilidade de ocorrência de um evento.⁵ O erro do nível 1 assumirá um determinado valor quando a variável dependente for igual a um, e outro valor quando ela for igual a zero (RIANI, 2005).

5 Neste trabalho, a estratégia empírica considera a estrutura hierárquica existente nos dados educacionais (por exemplo, alunos, turmas, escolas). Ao tratar variáveis em diferentes níveis, o modelo de regressão clássico pode não ser o mais apropriado, por não levar em conta a presença de correlação entre indivíduos relacionados em um mesmo nível de agregação (LAROS; MARCIANO, 2008).

No presente caso, por exemplo, esse problema seria representado pela correlação entre alunos de uma mesma escola. Assim, a abordagem multinível é preferível à linear por considerar os diversos níveis de agregação, corrigindo os erros-padrão, testes de hipóteses e intervalos de confiança, além do risco que há de se perder a importância dos efeitos de grupo e efeitos aleatórios dos parâmetros ao utilizar o enfoque tradicional.

Assim, no modelo logístico multinível padrão, as observações em y são aleatórias e independentes, com distribuição Bernoulli, estando condicionadas a um vetor de efeitos aleatórios. A modelagem usual para o caso de respostas binárias utiliza uma função *logit* (MCMAHON; POUGET; TORTU, 2006; FIELDING; GOLDSTEIN, 2006), conforme equação abaixo:

$$\text{logit}(p_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \eta_{ij} \quad (1)$$

Onde: p_{ij} é a probabilidade de ocorrência do evento, para o indivíduo i no grupo j ; e η_{ij} é o logaritmo das chances de ocorrência do evento.

Neste trabalho, optou-se por um modelo de dois níveis no contexto de atraso escolar. Têm-se, portanto, estudantes (i) que podem estar com dois anos ou mais além da idade adequada para a série ($y_{ij} = 1$) ou com idade regular ($y_{ij} = 0$), provenientes de escolas (j). O interesse está na probabilidade de o aluno se encontrar em atraso escolar. Destaca-se que essa probabilidade variará entre os alunos, dependendo dos valores de um conjunto de variáveis explicativas e do efeito da escola. Observa-se que o logaritmo das probabilidades de atraso é definido como a razão entre a probabilidade de atraso e a probabilidade de não atraso. O modelo logístico no nível 1 assume a seguinte forma:

$$\eta_{ij} = \beta_{0j} + \beta_{1j}x_{ij} \quad (2)$$

Onde: β_{0j} representa o valor de η_{ij} quando a variável independente assume o valor zero; e β_{1j} é a inclinação da η_{ij} em x_{ij} para a escola j .

Note-se que o termo de erro é omitido pelo modelo, visto que sua variância decorre diretamente da probabilidade de sucesso (segue uma distribuição binomial), a qual é assumida como fixa na transformação das chances de sucesso. Assim, não há um parâmetro de variância separado (FIELDING; GOLDSTEIN, 2006). Com isso, a formulação padrão do índice de correlação intraclasse pode não ser válida nesse caso (MCMAHON; POUGET; TORTU, 2006). Shirasu e Arraes (2015) sugerem utilizar o cálculo do Coeficiente de Partição da Variância (CPV), que é igual ao índice de correlação intraclasse em um modelo multinível simples, considerando a variância residual do nível 1 como $\sigma_{\epsilon}^2 = \frac{\pi^2}{3} = 3,29$ em virtude da modelagem logística adotada. Expandindo o modelo para o nível 2, tem-se que:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + u_{0j} \quad (3)$$

$$\beta_{1j} = \gamma_{10} \quad (4)$$

Em que: γ_{00} é o valor médio do intercepto; γ_{01} é a inclinação das variáveis explicativas z_j ; u_{0j} representa os efeitos aleatórios do nível 2; e γ_{10} corresponde ao efeito fixo entre os grupos.

Nesse caso, tem-se que, no primeiro nível, é definido um modelo *logit* usual para cada contexto (escola), permitindo que os coeficientes individuais variem ao longo das escolas. Substituindo as equações de nível 2 no modelo de nível 1, obtém-se a equação geral (5). Assim, nesse modelo, u_{0j} representa o único efeito aleatório (considerado na literatura o efeito-escola), sendo que os demais termos da equação são fixos.

$$\eta_{ij} = \gamma_{00} + \gamma_{01}z_j + \gamma_{10}x_{ij} + u_{0j} \quad (5)$$

Por fim, no uso desses modelos, comumente, utilizam-se transformações para se chegar a um estimador linear. E, no caso do modelo hierárquico, é a partir desse estimador linear que as estruturas multiníveis se manifestam por meio de efeitos aleatórios (FIELDING; GOLDSTEIN, 2006). Na maioria das

vezes, emprega-se a função inversa para o *logit* por meio da transformação exponencial (HOX, 2010).

VARIÁVEIS UTILIZADAS

Os quadros 1 e 2 apresentam a descrição das variáveis que entraram no modelo *logit*, conforme elenco reunido a partir da revisão da literatura e da disponibilidade na base de dados. Em resumo, as variáveis escolhidas para compor o modelo no nível do aluno (Quadro 1) foram gênero, cor, condição em relação ao trabalho, mora com os pais, escolaridade da mãe, tempo de lazer, realização de dever, nível socioeconômico (NSE) e nível de capital social (NCS) da família. No nível da escola (Quadro 2), as variáveis relacionadas ao contexto escolar para explicar o atraso escolar foram: esfera administrativa, localização, número de matrículas, índice de conservação da escola (ICE), índice de posse de equipamentos da escola (IPE), proficiência média, adoção de ciclo de progressão, forma de composição das turmas, programa de redução de abandono, proporção de professores com curso superior e macrorregião.⁶

⁶ As variáveis NSE, NCS, ICE e IPE são índices construídos a partir da Teoria da Resposta ao Item (TRI), um método que fornece indicadores capazes de discriminar melhor os indivíduos ao extrair mais informações dos dados. O escore obtido foi construído com base em um modelo de respostas graduadas, considerando uma série de itens dos questionários da Prova Brasil. O *software* usado para a elaboração dos escores foi o R Core Team (2017), sendo o processo de estimação o método de máxima verossimilhança.

Há diferenças na organização do ensino entre as redes municipais e estaduais. No ensino fundamental, algumas escolas estaduais adotam dois ciclos de alfabetização: um inicial, de três anos; e um ciclo complementar, com duração de dois anos. Cabe aos municípios a decisão sobre a forma de organização do sistema de ensino em relação à quantidade de ciclos e sua duração (RIANI; SILVA; SOARES, 2012). A inclusão da variável relativa aos ciclos decorre da adoção desse sistema que impede a retenção de alunos em determinadas séries (OLIVEIRA; SOARES, 2012), o que, provavelmente, tem relação com a probabilidade de o estudante estar em idade adequada para a série.

QUADRO 1 - Variáveis explicativas no nível do aluno

VARIÁVEL		DESCRIÇÃO
DEPENDENTE (RESPOSTA)		
Atraso escolar		Assume o valor "1" se o estudante tem dois anos ou mais de diferença da idade adequada para a série que frequenta e "0" para o caso contrário.
EXPLICATIVA NO NÍVEL DO ALUNO		
Homem		Assume o valor "1" se o aluno é do sexo masculino e "0" para o sexo feminino.
Branco		Igual a "1" se o aluno é branco e "0" para pardo, preto, amarelo, indígena e não sabe.
Mora com mãe e pai		Assume o valor "1" se o aluno mora com a mãe e o pai e "0" só com a mãe, só com o pai ou com outra pessoa responsável.
Escaridade da mãe (categoria base: não estudou)	4ª série/5º ano do EF completa/o	Assume o valor "1" se a mãe do aluno completou a 4ª série/5º ano, mas não completou a 8ª série/9º ano e "0" caso contrário.
	EF completo	Assume o valor "1" se a mãe do aluno completou a 8ª série/9º ano, mas não completou o ensino médio e "0" caso contrário.
	EM completo	Assume o valor "1" se a mãe do aluno completou o ensino médio, mas não completou a faculdade e "0" caso contrário.
	Superior completo	Assume o valor "1" se a mãe do aluno completou a faculdade e "0" caso contrário.
Trabalho doméstico		Assume o valor "1" se o aluno exerce algum trabalho doméstico e "0" caso contrário.
Trabalha fora		Assume o valor "1" se o aluno exerce algum trabalho fora de casa (recebendo ou não um salário) e "0" caso contrário.
Tempo de lazer		Número médio de horas gastas, em dias de aula, assistindo televisão, navegando na internet ou com jogos eletrônicos.
Dever de casa		Assume o valor "1" se o aluno faz o dever de casa sempre ou quase sempre e "0" caso contrário.
Nível socioeconômico do aluno (NSE)		Obtido pela aplicação da Teoria da Resposta ao Item (TRI) com respostas graduais, a partir dos seguintes itens do questionário do aluno: televisão; rádio; DVD; geladeira; geladeira duplex; freezer; máquina; carro; banheiro; empregada; escolaridade da mãe e do pai.
Nível de capital social do aluno (NCS)		Obtido pela aplicação da Teoria da Resposta ao Item (TRI) com respostas graduais, a partir dos seguintes itens do questionário: frequência com que os responsáveis vão às reuniões; pais incentivam a estudar; pais incentivam a fazer o dever de casa; pais incentivam a ler; pais incentivam a não faltar; se há diálogo sobre o que acontece na escola.

Fonte: Elaboração própria a partir de dados da Prova Brasil/Inep 2013.

Inclui-se no modelo uma variável que descreve o critério utilizado para a formação das turmas na escola, qual seja, se segue a homogeneidade ou heterogeneidade quanto à idade e ao rendimento. Um meio encontrado pelos gestores educacionais para tentar controlar variáveis do contexto escolar foi atuar na forma de composição das turmas, já que tal composição pode levar a diferenças em termos de rendimento. Espera-se, por exemplo, que a formação de turmas com estudantes de diferentes situações econômicas melhore o rendimento dos alunos menos privilegiados, mas há incertezas

quanto ao verdadeiro efeito no desempenho médio do grupo (SOARES, 2002). Foram incluídas, também, as macrorregiões do estado como variáveis de controle das diferenças culturais e regionais (PONTILI; KASSOUF, 2007).

QUADRO 2 - Variáveis explicativas no nível da escola

VARIÁVEL	DESCRIÇÃO	
EXPLICATIVA NO NÍVEL DA ESCOLA		
Municipal	Igual a "1" se a escola pertence à rede municipal e "0" caso contrário.	
Urbana	Igual a "1" se a escola está localizada na área urbana e "0" para rural.	
Matrículas 5º ano	Número de alunos matriculados no 5º ano, conforme o Censo 2013.	
Índice de posse de equipamentos da escola (IPE)	Obtido pela aplicação da TRI com respostas graduais, a partir dos seguintes itens: computador, copiadora, impressora, projetor de <i>slides</i> e televisão.	
Índice de conservação da escola (ICE)	Obtido pela aplicação da TRI com respostas graduais, a partir dos seguintes itens: telhado; paredes; piso; entradas do prédio; pátio; corredores; salas de aula; portas; janelas; banheiros; cozinha; instalações hidráulicas; instalações elétricas.	
Proficiência média	Proficiência média da escola obtida na prova de Português e Matemática.	
Adota ciclo [1]	Assume o valor "1" se a escola organiza o ensino fundamental em ciclos e "0" caso contrário.	
Formação de turmas heterogêneas	Assume o valor "1" se a escola, no processo de formação de turmas, opta pela heterogeneidade quanto à idade e rendimento e "0" caso contrário.	
Há programa de redução de abandono	Assume o valor "1" se a escola adota algum programa para reduzir taxas de abandono/evasão e "0" caso contrário.	
Proporção de professores com curso superior	Proporção de professores com curso superior na escola.	
Macrorregiões (categoria base: norte).	Alto Paranaíba	Assume o valor "1" se a escola está localizada na região e "0" caso contrário (categoria base: norte).
	Central	
	Centro-Oeste	
	Jequitinhonha e Mucuri	
	Zona da Mata	
	Noroeste	
	Rio Doce	
	Sul	
Triângulo		

Fonte: Elaboração própria a partir de dados da Prova Brasil/Inep 2013.

[1] As informações para esta variável foram extraídas dos microdados do Censo Escolar de 2013.

As variáveis foram centradas na grande média. Para o modelo logístico multinível, essa transformação tem como vantagem a estimativa da probabilidade de atraso prevista

no valor médio das variáveis independentes somente pelo intercepto, visto que, na média, os valores das variáveis explicativas são iguais a zero (RIOS-NETO; CEZAR; RIANI, 2002).

ESTATÍSTICAS DESCRITIVAS

As tabelas 2 a 5 trazem as estatísticas descritivas – média, desvio-padrão, valor máximo e mínimo e coeficiente de variação (CV) das características dos alunos e das escolas que compõem a amostra final no 5º e 9º ano,⁷ para o estado de Minas Gerais. Para não tornar exaustiva a leitura, optou-se pela discussão apenas das informações para o 9º ano.

De acordo com os dados da Tabela 3, em média, 16% dos estudantes do 9º ano estavam com idade diferente da recomendada para a série, em 2013. Observa-se que cerca de 47% dos alunos são meninos e, em relação à cor, aproximadamente 31% dos estudantes se declararam brancos. No que diz respeito à escolaridade materna, em média, 26% tinham até a 4ª série do fundamental completa, 17% tinham ensino fundamental completo, 26% concluíram o ensino médio e 12% tinham superior completo. Nota-se que 69% dos alunos moravam com os pais e grande parte deles (82%, em média) realizava tarefas domésticas e 19% trabalhavam fora de casa.

O tempo gasto pelos alunos, em média, com TV, internet ou jogos eletrônicos, para a amostra, foi de 2,73 horas, com desvio-padrão de 1,24 hora. Na amostra, 61% dos alunos afirmaram que faziam as tarefas, tanto de Português como de Matemática, na maior parte das vezes. Em média, os estudantes possuíam um NSE de 0,04, com um coeficiente de variação de 23,37, indicando uma elevada variabilidade dos dados em relação à média e uma alta heterogeneidade da condição socioeconômica dos discentes. A média do NCS foi de -0,08, com coeficiente de variação de 7,81, mostrando bastante heterogeneidade quanto ao nível de participação e atenção dos pais na vida escolar dos seus filhos.

⁷ Os dados da Prova Brasil 2013 foram filtrados para os alunos de Minas Gerais, matriculados no 5º ano do ensino fundamental (ou no 9º ano para o segundo caso), que preencheram a prova, preencheram o questionário, estavam em concordância com o Censo Escolar e que pertenciam à rede municipal ou estadual de ensino.

TABELA 2 - Estatísticas descritivas no nível do aluno, 5º ano, Minas Gerais (1)

VARIÁVEL		5º ANO DO ENSINO FUNDAMENTAL				
		MÉDIA	D. PADRÃO	MÍNIMO	MÁXIMO	CV
Atraso escolar		0,11	-	0	1	-
Homem		0,50	-	0	1	-
Branco		0,30	-	0	1	-
Mora com mãe e pai		0,71	-	0	1	-
Escolaridade da mãe (categoria base: não estudou)	4ª série do ensino fundamental	0,22	-	0	1	-
	Ensino fundamental completo	0,17	-	0	1	-
	Ensino médio completo	0,24	-	0	1	-
	Superior completo	0,20	-	0	1	-
Trabalho doméstico		0,80	-	0	1	-
Trabalho fora		0,12	-	0	1	-
Tempo de lazer		2,34	1,24	0	4	0,53
Dever de Matemática		0,85	-	0	1	-
Dever de Português		0,83	-	0	1	-
Nível socioeconômico		0,31	0,87	-3,18	4,43	2,81
Nível de capital social		-0,04	0,57	-2,76	0,32	-14,3
Observações (alunos do 5º ano do EF)		35.787				

Fonte: Elaboração própria a partir de dados da Prova Brasil/Inep 2013.

(1) Para as variáveis binárias entende-se a média como a proporção de sucesso, isto é, a proporção da variável *dummy* quando ela é igual a 1.

TABELA 3 - Estatísticas descritivas no nível do aluno, 9º ano, Minas Gerais

VARIÁVEL		9º ANO DO ENSINO FUNDAMENTAL				
		MÉDIA	D. PADRÃO	MÍNIMO	MÁXIMO	CV
Atraso escolar		0,16	-	0	1	-
Homem		0,47	-	0	1	-
Branco		0,31	-	0	1	-
Mora com mãe e pai		0,69	-	0	1	-
Escolaridade da mãe (categoria base: não estudou)	4ª série do ensino fundamental	0,26	-	0	1	-
	Ensino fundamental completo	0,17	-	0	1	-
	Ensino médio completo	0,26	-	0	1	-
	Superior completo	0,12	-	0	1	-
Trabalho doméstico		0,82	-	0	1	-
Trabalho fora		0,19	-	0	1	-
Tempo de lazer		2,73	1,24	0	4	0,45
Dever de Matemática		0,61	-	0	1	-
Dever de Português		0,61	-	0	1	-
Nível socioeconômico		0,04	0,88	-3,73	5,05	23,37
Nível de capital social		-0,08	0,66	-2,80	0,43	-7,81
Observações (alunos do 9º ano do EF)		54.624				

Fonte: Elaboração própria a partir de dados da Prova Brasil/Inep 2013.

TABELA 4 - Estatísticas descritivas no nível da escola, 5º ano, Minas Gerais

VARIÁVEL		ESCOLAS COM 5º ANO DO ENSINO FUNDAMENTAL				
		MÉDIA	D. PADRÃO	MÍNIMO	MÁXIMO	CV
Municipal		0,63	-	0	1	-
Urbana		0,93	-	0	1	-
Matrículas		62,88	34,73	20	269	0,55
Índice de posse de equipamentos da escola		0,06	0,73	-3,1	0,78	11,63
Índice de conservação da escola		-0,07	0,83	-2,8	0,97	-11,96
Proficiência média		223,75	21,84	155,21	314,97	0,1
Adota ciclo		0,81	-	0	1	-
Formação de turmas heterogêneas		0,46	-	0	1	-
Há programa de redução de abandono		0,71	-	0	1	-
Proporção de professores com curso superior		0,94	-	0	1	-
Macrorregiões (categoria base: norte)	Alto Paranaíba	0,04	-	0	1	-
	Central	0,31	-	0	1	-
	Centro-Oeste	0,06	-	0	1	-
	Jequitinhonha e Mucuri	0,06	-	0	1	-
	Zona da Mata	0,12	-	0	1	-
	Noroeste	0,02	-	0	1	-
	Norte	0,08	-	0	1	-
	Rio Doce	0,10	-	0	1	-
	Sul	0,13	-	0	1	-
	Triângulo	0,07	-	0	1	-
Observações (escolas com 5º ano do EF)		2.278				

Fonte: Elaboração própria a partir de dados da Prova Brasil/Inep 2013.

A Tabela 5 traz as estatísticas descritivas das variáveis relativas à escola para o 9º ano. Na amostra, 33% das escolas são da rede municipal e a maior parte está localizada na área urbana (93%). A média de matrículas no 9º ano foi de 70,55 alunos, mas as escolas diferem quanto a esse aspecto (coeficiente de variação de 0,61).

No que diz respeito ao ICE, observam-se média de -0,19 e desvio-padrão de 0,83, sugerindo que há uma volatilidade das condições de conservação da infraestrutura entre as escolas. O IPE alcançou uma média de -0,003 e um elevado coeficiente de variação de 235,25, sinalizando a presença de heterogeneidade entre as escolas quanto à condição de conservação dos equipamentos. O valor médio da proficiência média de Português e Matemática foi de 259,56 pontos, com desvio-padrão de 18,50

pontos. O baixo valor encontrado para o coeficiente de variação (0,07) sugere que as escolas são homogêneas em termos da proficiência média obtida pelos seus estudantes nas duas disciplinas.

TABELA 5 - Estatísticas descritivas no nível da escola, 9º ano, Minas Gerais

VARIÁVEL		ESCOLAS COM 9º ANO DO ENSINO FUNDAMENTAL				
		MÉDIA	D. PADRÃO	MÍNIMO	MÁXIMO	CV
Municipal		0,33	-	0	1	-
Urbana		0,93	-	0	1	-
Matrículas		70,55	43,07	20	356	0,61
Índice de posse de equipamentos da escola		0,00	0,74	-3,10	0,78	-235,25
Índice de conservação da escola		-0,19	0,83	-2,80	0,97	-4,46
Proficiência média		259,56	18,50	192,50	317,67	0,07
Adota ciclo		0,88	-	0	1	-
Formação de turmas heterogêneas		0,43	-	0	1	-
Há programa de redução de abandono		0,89	-	0	1	-
Proporção de professores com curso superior		0,97	-	0	1	-
Macrorregiões	Alto Paranaíba	0,04	-	0	1	-
	Central	0,31	-	0	1	-
	Centro-Oeste	0,05	-	0	1	-
	Jequitinhonha e Mucuri	0,06	-	0	1	-
	Zona da Mata	0,12	-	0	1	-
	Noroeste	0,02	-	0	1	-
	Norte	0,10	-	0	1	-
	Rio Doce	0,10	-	0	1	-
	Sul	0,13	-	0	1	-
Triângulo	0,07	-	0	1	-	
Observações (escolas com o 9º ano do EF)		1.793				

Fonte: Elaboração própria a partir de dados da Prova Brasil/Inep 2013.

Na amostra, em média, 88% das escolas organizam o ensino fundamental em ciclos. Nota-se que, para a composição das turmas, em média, 43% das escolas adotam a opção pela heterogeneidade quanto à idade e ao rendimento. De outro lado, grande parte das escolas (89%) tem algum programa de redução de taxas de abandono. Observa-se, ainda, que a maioria dos professores possui ensino superior (97%).

Por fim, em termos regionais, percebe-se que a maior parte das escolas localiza-se na região central do estado,

abrangendo a Região Metropolitana de Belo Horizonte, que reúne o maior número de municípios.

RESULTADOS DO MODELO HIERÁRQUICO PARA O ATRASO ESCOLAR

A Tabela 6 mostra os resultados do modelo nulo para a probabilidade de atraso escolar em Minas Gerais. Foram estimados modelos para os dados da Prova Brasil, dos 5º e 9º anos do ensino fundamental, na tentativa de identificar se existem diferenças entre os fatores associados ao atraso escolar nos anos iniciais e finais do ensino fundamental. O programa utilizado nas estimações foi o Stata 14.0.

TABELA 6 – Modelo nulo

VARIÁVEIS EXPLICATIVAS	5º EF			9º EF		
	ESTIMATIVA	EP	P-VALOR	ESTIMATIVA	EP	P-VALOR
EFEITO FIXO						
Intercepto (γ_{00})	-2,33	0,003	0,000	-1,71	0,02	0,000
EFEITO ALEATÓRIO	VARIÂNCIA	EP	RAZÃO-T	VARIÂNCIA	EP	RAZÃO-T
Variância do intercepto (σ_{u0}^2)	0,787	0,06	13,1	0,42	0,02	21,0
CPV	0,1929			0,1132		
Deviance M1	23.878,49			47.074,31		
Número de parâmetros	2			2		
Número de observações	35.787			54.624		
Número de grupos	2.278			1.793		

Fonte: Elaboração própria a partir de dados da Prova Brasil/Inep 2013.

Ao exponenciar a estimativa do coeficiente de regressão, tem-se a razão de chances (*odds ratio* – OR), que representa como as chances de atraso se modificam quando se passa entre as diferentes categorias de um fator. Assim, para cada coeficiente das variáveis explicativas na regressão logística, a OR indica a diferença percentual nas chances de estar em atraso quando a variável independente aumenta em uma unidade, mantendo-se as demais variáveis constantes (ORTIGÃO; AGUIAR, 2013).

Quando a razão de chance for inferior a 1, a variável explicativa reduz a probabilidade de o evento ocorrer e, quando for superior a 1, aumenta-se a possibilidade de ocorrência

do evento, ou seja, essa razão quantifica como cada variável explicativa relaciona-se com as probabilidades de cada valor de resposta da variável dependente.

Interpretando os resultados do modelo nulo (Tabela 6), verifica-se que a estimativa do intercepto (na terceira linha/segunda e quinta colunas) indica que as chances de o aluno ter dois anos ou mais da idade adequada em relação à média das escolas ($u_{0j} = 0$) são de 0,10, no 5º ano, e de 0,18, no 9º ano. A variância do intercepto em log de chances do atraso é 0,787, com erro padrão de 0,06. Segundo o teste t, essa variância é estatisticamente significativa, mostrando que há diferenças entre as escolas em termos da probabilidade de atraso. Já o coeficiente de partição da variância (CPV) revela que cerca de 19% e 11% da variabilidade na probabilidade de atraso escolar ocorre devido às características escolares, respectivamente, no 5º e 9º anos, demonstrando a necessidade da modelagem hierárquica.

A Tabela 7 apresenta os resultados do modelo logístico hierárquico para a probabilidade de atraso escolar. Observa-se que, para a variável que indica o sexo do aluno, encontrou-se a razão de chance para o estudante que se declarou homem igual a 1,74. Esse valor indica que a chance de os meninos estarem em atraso é 74% maior do que as meninas, no caso do 5º ano do ensino fundamental, mantidas constantes as demais variáveis. Para o 9º ano, a chance de estar em atraso é 60% maior para os meninos. Esse resultado está em consonância com a literatura, que tem encontrado diferenças em termos de gênero na trajetória escolar.

A probabilidade de atraso entre os alunos que se declararam brancos é menor do que para aqueles que se declararam pretos, pardos, amarelos e índios, em ambas as séries. Nota-se que a razão de chances de atraso para brancos é 16% menor do que para não brancos, no 9º ano, e 23%, no 5º ano, refletindo a presença de desigualdades associadas à cor nas escolas.

Os resultados do modelo sugerem que alunos que moram com uma família tradicional, formada por pai e mãe, têm probabilidade menor de estarem em atraso escolar. Nos anos iniciais do ensino fundamental, por exemplo, a chance

de os alunos que residem com pai e mãe apresentarem atraso é 44% menor do que a dos estudantes que não moram com o pai e a mãe. Espera-se, portanto, que famílias formadas por pai, mãe e filhos exibam condições de socialização mais favoráveis aos jovens (SILVA; HASENBALG, 2002).

A educação da mãe, tida como uma referência do capital cultural familiar, é um fator importante para que os estudantes frequentem a escola na idade correta. Tomando-se como referência a categoria-base “mãe não estudou”, observa-se que, quanto mais escolarizada é a mãe, menor se torna a probabilidade de atraso escolar do filho, sendo esse efeito maior para o 5º do que para o 9º ano, o que pode ser indicativo da maior influência materna nas decisões escolares das crianças nos anos iniciais de ensino.

TABELA 7 – Modelo logístico hierárquico para a probabilidade de atraso

VARIÁVEIS EXPLICATIVAS		5º EF			9º EF		
EFEITO FIXO		ESTIMATIVA	P-VALOR	O.R.	ESTIMATIVA	P-VALOR	O.R.
Intercepto		-2,630	0,000	0,07	-1,947	0,000	0,14
ALUNO							
Homem		0,056	0,000	1,74	0,468	0,000	1,60
Branco		-0,026	0,000	0,77	-0,178	0,000	0,84
Mora com mãe e pai		-0,059	0,000	0,56	-0,442	0,000	0,64
Escolaridade da mãe (categoria base= não estudou)	Até 4ª série do EF	-0,025	0,000	0,78	-0,516	0,000	0,60
	EF completo	-0,050	0,000	0,61	-0,719	0,000	0,49
	EM completo	-1,125	0,000	0,32	-0,976	0,000	0,38
	Superior completo	-1,153	0,000	0,32	-1,140	0,000	0,32
Trabalho doméstico		0,002	0,619	1,02	-0,070	0,036	0,93
Trabalho fora de casa		0,071	0,000	2,03	0,638	0,000	1,89
Tempo de lazer		-0,005	0,001	0,95	-0,173	0,000	0,84
Dever de Matemática		-0,028	0,000	0,75	-0,322	0,000	0,72
Dever de Português		-0,024	0,000	0,79	-0,165	0,000	0,85
NSE		-0,006	0,048	0,94	-0,022	0,240	0,98
NCS		-0,020	0,000	0,82	-0,155	0,000	0,86
ESCOLA							
Municipal		0,064	0,000	1,89	-0,130	0,009	0,88
Urbana		-0,007	0,596	0,94	-0,181	0,039	0,83
Matriculas		0,000	0,973	1,00	-0,001	0,182	1,00
ICE		0,001	0,742	1,01	0,026	0,302	1,03
IPE		-0,001	0,892	0,99	-0,029	0,312	0,97
Proficiência média		-0,001	0,000	0,99	-0,009	0,000	0,99
Adota ciclo		-0,040	0,000	0,67	-0,322	0,000	0,72
Formação de turmas heterogêneas		0,011	0,027	1,12	0,147	0,000	1,16
Há programa de redução de abandono		0,013	0,027	1,13	-0,018	0,778	0,98
Prop. Professor com curso superior		0,002	0,832	1,02	-0,018	0,885	0,98
Macrorregiões (categoria base= Norte)	Alto Paranaíba	0,067	0,000	1,95	0,282	0,014	1,33
	Central	0,041	0,000	1,51	0,550	0,000	1,73
	Centro-Oeste	0,035	0,024	1,42	0,170	0,160	1,18
	Jequitinhonha e Mucuri	0,058	0,000	1,79	0,489	0,000	1,63
	Zona da Mata	0,078	0,000	2,18	0,611	0,000	1,84
	Noroeste	0,034	0,093	1,40	0,028	0,863	1,03
	Rio Doce	0,037	0,004	1,45	0,482	0,000	1,62
	Sul	0,065	0,000	1,91	0,383	0,000	1,47
	Triângulo	0,063	0,000	1,89	0,282	0,006	1,33
EFEITO ALEATÓRIO		VARIÂNCIA	E. P.	RAZÃO-T	VARIÂNCIA	E. P.	RAZÃO-T
Intercepto (σ_{u0}^2)		0,383	0,039	9,94	0,308	0,023	13,19
NSE		0,082	0,028	2,93	0,046	0,013	3,59
CPV		0,106			0,086		

Fonte: Elaboração própria a partir de dados da Prova Brasil/Inep 2013.

Os resultados do modelo para o 5º ano mostram que, em comparação aos alunos com mães que não estudaram ou não completaram o primeiro ciclo do ensino fundamental (categoria-base), a probabilidade de atraso é, em média: 22% menor para estudantes com mães que cursaram até a 4ª série; 39% menor para aqueles com mães que completaram o ensino fundamental; e 68% menor para alunos com mães que concluíram o ensino médio e para aqueles com mães que possuem curso superior completo. Rios-Neto, Cezar e Riani (2002), em suas conclusões, sugerem que, nas séries iniciais e nos locais com maior proporção de mães com baixa escolaridade, deveriam ser alocados os melhores professores com a finalidade de melhorar a equidade da educação nessas regiões.

A variável relativa ao trabalho doméstico não teve significância estatística para alunos do 5º ano, mas foi significativa para os do 9º ano. Nesse último caso, observa-se que as chances de atraso são 7% menores para alunos que realizam tarefas domésticas. Para o 5º ano, o trabalho fora de casa, remunerado ou não, aumenta a probabilidade de atraso escolar em duas vezes mais do que para os alunos que não trabalham fora. Para o 9º ano, esse valor é estimado em 1,89. Desse modo, estudantes que trabalham fora de casa têm menor chance de avançar nos estudos se comparados aos que não trabalham (LEON; MENEZES-FILHO, 2002).

O tempo despendido em lazer, seja assistindo à televisão, navegando na internet ou com jogos eletrônicos, reduz as chances de atraso em 5%, no 5º ano, e 16%, no 9º ano. A realização de dever de casa pelo aluno é vista como um indicador do interesse ou motivação pelos estudos. Por isso, incluiu-se no modelo a variável que capta se o aluno realiza ou não as lições de casa, nas duas disciplinas avaliadas na Prova Brasil. Como esperado, os estudantes que afirmaram fazer dever de casa têm chances menores de estarem em atraso escolar. Em relação à Matemática, as chances de atraso escolar do aluno que afirmou fazer o dever de casa são 25% (28%) menores no 5º ano (9º ano). No caso de Língua Portuguesa, as chances de atraso escolar são 21% (15%) menores no 5º ano (9º ano), em comparação com os estudantes que não fazem as tarefas. Portanto, o empenho em realizar o dever indica a

responsabilidade do discente com o aprendizado, o que pode reduzir a probabilidade de ele reprovar ou abandonar a escola, entre outros fatores que levam ao atraso nos estudos.

O NSE do aluno não foi estatisticamente significativo para explicar a probabilidade de atraso escolar em ambas as séries. Essa perda de significância está relacionada à inclusão do controle do contexto escolar no modelo. Por outro lado, o NCS, que representa a importância do interesse dos pais pela vida escolar de seus filhos, foi significativo, indicando que jovens com um NCS acima da média têm menores chances de ter dois anos a mais de diferença da idade adequada para a série que cursa. No 5º ano do ensino fundamental, a chance de estar em atraso é 18% menor para os estudantes com NCS acima da média escolar em comparação àqueles com um capital social familiar abaixo da média. Para o 9º ano, nesse caso, a chance de atraso foi 14% menor. Assim, a trajetória escolar está relacionada de forma direta com a postura dos pais que vão às reuniões, incentivam os estudos, a leitura, a realização de dever, marcam presença nas aulas e conversam sobre a escola. Esses resultados confirmam achados de outros estudos que analisaram a importância de um ambiente familiar favorável à educação como determinante no desempenho escolar (ORTIGÃO; AGUIAR, 2013; PALERMO; SILVA; NOVELLINO, 2014).

Passando à análise das variáveis relacionadas ao ambiente escolar, observam-se diferenças na chance de estar em atraso escolar, segundo a esfera administrativa da escola e as séries. Para o 5º ano, a probabilidade de um aluno de escola municipal estar atrasado é 89% maior, quando comparado a um discente de escola estadual. Já para o 9º ano, a chance de um aluno do ensino municipal estar em idade superior à adequada para a série é 12% menor, quando comparado aos estudantes da rede estadual.

Como esperado, a adoção de ciclos de progressão pela escola diminui a probabilidade de atraso escolar, visto que esse sistema impede a reprovação em determinadas séries. Para alunos do 5º ano que estudam em escolas que adotam a política de ciclos, a chance de estar em atraso escolar é reduzida em 33%. Já para os estudantes do 9º ano, a probabilidade

de atraso escolar é 28% menor nas escolas que adotam ciclos, quando comparadas àquelas que não adotam.

A variável que se refere à localização da escola buscou captar os diferenciais de atraso escolar entre escolas de zona urbana e rural, não sendo essa variável estatisticamente significativa para o 5º ano. Nota-se que, para o último ano do ensino fundamental, a localização da escola na área urbana reduz em 17% a chance de atraso escolar, em comparação com as escolas rurais. Isso pode estar relacionado a dificuldades de transporte, à precariedade da infraestrutura da escola, à carência de professores nessas áreas, ao trabalho dos jovens (custo de oportunidade de continuar estudando) na propriedade e, também, à falta de políticas educacionais específicas para a área rural.

A variável matrículas foi incluída para tentar captar o tamanho da escola e, com isso, controlar a heterogeneidade escolar. Para cada ano, utilizou-se o número de alunos matriculados na série. No entanto, no presente modelo, essa variável não apresentou significância estatística. Do mesmo modo, os indicadores de conservação (ICE) e de posse de equipamentos (IPE) também não foram estatisticamente significativos para explicar o atraso escolar.

Honda (2006) aponta que os dados do Saeb revelam que estudantes com defasagem idade-série apresentavam proficiências médias menores do que os estudantes sem atraso escolar, evidenciando a forte relação entre as duas variáveis. No trabalho de Paula (2017) também encontraram-se evidências de que os alunos em atraso escolar (tanto em Minas Gerais como no Ceará) têm seu desempenho penalizado na Prova Brasil. No caso cearense, nota-se ainda que a influência da diferença de dois anos ou mais na idade considerada adequada para a série é diferente entre as escolas. Assim, tentando verificar se há relação entre o desempenho individual do aluno e a probabilidade de ele estar com idade diferente da recomendada para a série, inclui-se no modelo a nota média do aluno obtida nas provas de Matemática e Língua Portuguesa. Em ambas as séries, a razão de chances encontrada foi de 0,99, indicando que cada ponto adicional na proficiência média do estudante reduz a chance de

atraso, em média, em 1%. Assim, políticas que levem a melhorias no desempenho, principalmente, dos estudantes com menores notas podem ter impactos significativos na redução das taxas de atraso escolar.

A forma de composição das turmas foi incluída também para controle do contexto escolar. Observou-se que escolas que adotam o critério da heterogeneidade (quanto a idade e rendimento) na formação de turmas têm chances de atraso escolar maiores em 12% no 5º ano e 16% no 9º ano em relação àquelas que optam pela homogeneidade. Nesse caso, a influência dos pares na trajetória escolar parece ter consequências negativas, sinalizando que as escolas devem estar atentas para que a forma de composição das turmas não cause elevadas diferenças no rendimento.

Outra questão analisada foi em relação às ações da escola para reduzir a taxa de abandono. Os resultados mostraram que, nas escolas que adotam programas de redução das taxas de abandono, as chances de atraso escolar são 13% maiores do que naquelas que não adotam ou não apresentam esse problema para o 5º ano (no 9º ano, não houve significância estatística). O que pode explicar essa associação é que, possivelmente, tais ações levam à retenção de alunos com baixo desempenho que estão em atraso, ou porque onde se implementam essas ações já persista uma taxa muito elevada de evasão escolar.

Também foi incluída no modelo a proporção de professores com curso superior, com a finalidade de avaliar se a maior escolaridade do docente está relacionada com o atraso escolar. No entanto, essa variável não apresentou significância estatística, talvez porque grande parte da amostra seja composta por professores com ensino superior (97%).

Por fim, as *dummies* para região buscaram captar as diferenças regionais na probabilidade de atraso. A região norte de Minas Gerais foi escolhida como a categoria de base, pois apresentou a menor proporção de alunos com atraso escolar no banco de dados. No geral, observam-se diferenças regionais em termos da probabilidade de atraso escolar. Por exemplo, para o 5º ano, em comparação aos alunos residentes no norte de Minas, a probabilidade de estar em atraso

é superior em 1,95 vez para alunos do Alto Paranaíba, 1,51 vez para aqueles da região central do estado, 1,42 vez para os do centro-oeste, 1,79 vez para alunos de escolas do Vale do Jequitinhonha e Mucuri, 2,18 vezes para os residentes na Zona da Mata, 1,45 vez para discentes do Rio Doce, 1,91 vez para alunos do Sul de Minas e 1,89 vez para aqueles do Triângulo Mineiro. Não teve significância estatística a *dummy* para a região noroeste do estado, em ambas as séries, e para o centro-oeste, no 9º ano.

Passando-se à análise dos efeitos aleatórios do modelo, ao considerar que o intercepto varia entre as escolas, assume-se que há um efeito diferenciado sobre a probabilidade do atraso escolar dependendo das características da escola. Dos resultados do modelo, pode-se considerar que o intercepto e o nível socioeconômico do estudante tenham efeitos aleatórios. Comparando as duas séries, nota-se que o efeito aleatório é maior para o 5º ano do ensino fundamental, tanto no intercepto como para o NSE. Portanto, a probabilidade média de um aluno estar em atraso escolar é diferenciada entre as escolas, assim como o NSE do aluno relaciona-se diferentemente com a probabilidade de ele frequentar a série na idade correta. Em outras palavras, os componentes de variância resumem, para cada estabelecimento de ensino, o modo com que a probabilidade de o aluno estar em atraso difere sistematicamente dos estudantes (com características similares) matriculados em outras escolas.

Da construção de intervalos de confiança, por exemplo, no caso do 5º ano, para o intercepto, tem-se que as chances de o aluno estar em atraso escolar em relação à média escolar são de 0,02, no limite inferior, e 0,24, no limite superior. Para o NSE, o intervalo de confiança mostra que cada ponto adicional no NSE acima da média reduz as chances de atraso em 43%, no limite inferior, e aumenta as chances de atraso em 74%, no limite superior.

Após realizado o controle do modelo por todas essas características apresentadas, o cálculo do CPV para o modelo final aponta que cerca de 11% e 9% da variabilidade na probabilidade de atraso escolar é explicada por fatores escolares, respectivamente, no 5º e 9º anos. Esse resultado sugere que

políticas de intervenção nas escolas podem ser importantes para reduzir as taxas de atraso escolar.

CONSIDERAÇÕES FINAIS

Este artigo buscou aprofundar o conhecimento a respeito dos motivadores do atraso escolar nos anos iniciais e finais do ensino fundamental. A partir da análise da relação entre as características dos estudantes e de fatores escolares sobre o atraso, notou-se que têm maiores chances de atraso alunos do sexo masculino, não brancos, que trabalham fora de casa, estudam em turmas heterogêneas e em escolas que adotam programas de redução de taxas de abandono.

Em contraste, as chances de estar em atraso escolar são menores para os estudantes que moram com os pais, têm mães mais escolarizadas, dispõem de tempo de lazer, fazem o dever de casa, têm pais que apoiam os estudos, estudam em escolas que adotam ciclos no ensino fundamental e naquelas que possuem níveis maiores na proficiência média em Matemática e Língua Portuguesa. Além disso, há diferenças entre as regiões na probabilidade de atraso escolar.

Em termos de política, é importante oferecer uma visão sobre como analisar diferentes propostas para lidar com elevadas taxas de atraso escolar. Nota-se que o trabalho infantil duplica as chances de atraso escolar em comparação aos estudantes que não trabalham, mostrando quão prejudicial isso pode ser para a vida escolar desses jovens. Em relação à opção pela adoção de ciclos ou sistemas de progressão continuada pelas escolas, observou-se que ela reduz a probabilidade de atraso escolar, já que esse sistema evita a reprovação em algumas séries. No entanto, é preciso estar atento a que custos essas ações são positivas para a produção de um aprendizado de qualidade.

A utilização dos modelos multinível permitiu a análise do relacionamento mútuo do nível familiar e do nível escolar na probabilidade de o aluno estar em atraso escolar, sendo que as escolas constituem o lócus das intervenções por parte das políticas educacionais (RIOS-NETO; CEZAR; RIANI, 2002). Assim, espera-se que a presente análise contribua

para a discussão das causas do atraso escolar, destacando-se a relevância desses resultados para a melhoria da qualidade da educação pública em Minas Gerais, que ainda precisa superar alguns problemas persistentes, tais como as elevadas taxas de distorção idade-série e o abandono. Portanto, sugere-se que políticas destinadas à redução das taxas de atraso escolar, as quais levem em conta a importância dos fatores aqui apresentados (por exemplo, a escolaridade da mãe, o NCS da família, o trabalho infantil, o NSE, a composição das turmas, a forma de organização do ensino, entre outros), poderão contribuir para o entendimento e o combate aos fatores que provocam o atraso escolar no estado de Minas Gerais.

REFERÊNCIAS

BIONDI, R. L.; FELÍCIO, F. *Atributos escolares e o desempenho dos estudantes: uma análise em painel dos dados Saeb*. Brasília, DF: Inep, 2007.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Prova Brasil*. Microdados, 2013.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Taxas de distorção idade-série*. 2015. Disponível em: <http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2015/distorcao_idade_serie/tdi_brasil_regioes_UFs_2015.zip>. Acesso em: 16 jan. 2017.

CURI, A. Z.; MENEZES-FILHO, N. A. A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil. *Estudos Econômicos*, São Paulo, v. 39, n. 4, p. 811-850, out./dez. 2009.

FERRÃO, M. E.; BELTRÃO, K. I.; SANTOS, D. P. dos. O impacto de políticas de não-repetência sobre o aprendizado dos alunos da 4ª série. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 32, n. 3, p. 495-513, dez. 2002.

FIELDING, A.; GOLDSTEIN, H. *Cross-classified and multiple membership structures in multilevel models: an introduction and review*. Birmingham: University of Birmingham, 2006. (Research Report, n. 791).

GOMES-NETO, J. B.; HANUSHEK, E. A. Causes and consequences of grade repetition: evidence from Brazil. *Economic Development and Cultural Change*, Chicago, v. 43, n. 1, p. 117-148, Oct. 1994.

HONDA, K. M. *Um estudo sobre os determinantes do atraso escolar*. 2006. 80 f. Dissertação (Mestrado em Economia Aplicada) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, Ribeirão Preto, 2006.

HOX, J. *Multilevel analysis: techniques and applications*. Mahwah: Lawrence Erlbaum Associates, 2010.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Síntese de indicadores sociais: uma análise das condições de vida da população brasileira*. Rio de Janeiro: IBGE, 2016. (Estudos & Pesquisas – Informação Demográfica e Socioeconômica, n. 36).

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. Educação. *Políticas Sociais – Acompanhamento e Análise*, Brasília, DF, n. 22, p. 225-285, set. 2014.

LAROS, J. A.; MARCIANO, J. L. P. Análise multinível aplicada aos dados do NELS: 88. *Estudos em Avaliação Educacional*, São Paulo, v. 19, n. 40, p. 263-278, maio/ago. 2008.

LEON, F. L. L.; MENEZES-FILHO, N. A. Reprovação, avanço e evasão escolar no Brasil. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 32, n. 3, p. 417-451, dez. 2002.

MCMAHON, J. M.; POUGET, E. R.; TORTU, S. A guide for multilevel modeling of dyadic data with binary outcomes using SAS PROC NLMIXED. *Computational Statistics & Data Analysis*, v. 50, n. 12, p. 3663-3680, Aug. 2006.

OLIVEIRA, L. F. B.; SOARES, S. *Determinantes da repetência escolar no Brasil: uma análise do painel dos censos escolares entre 2007 e 2010*. Brasília, DF: Ipea, 2012. (Texto para discussão, 1706).

ORTIGÃO, M. I. R.; AGUIAR, G. S. Repetência escolar nos anos iniciais do ensino fundamental: evidências a partir dos dados da Prova Brasil 2009. *Revista Brasileira de Estudos Pedagógicos*, Brasília, DF, v. 94, n. 237, p. 364-389, maio/ago. 2013.

PALERMO, G. A.; SILVA, D. B. N.; NOVELLINO, M. S. F. Fatores associados ao desempenho escolar: uma análise da proficiência em Matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro. *Revista Brasileira de Estudos de População*, Rio de Janeiro, v. 31, n. 2, p. 367-394, jul./dez. 2014.

PAULA, J. S. *Determinantes do desempenho educacional no Estado de Minas Gerais*. 2017. 157 f. Tese (Doutorado em Economia) – Instituto de Economia, Universidade Federal de Uberlândia, Uberlândia, 2017.

PONTILI, R. M.; KASSOUF, A. L. Fatores que afetam a frequência e o atraso escolar, nos meios urbano e rural, de São Paulo e Pernambuco. *Revista de Economia e Sociologia Rural*, Brasília, DF, v. 45, n. 1, p. 27-47, jan./mar. 2007.

R CORE TEAM. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2017. Disponível em: <<https://www.R-project.org/>>. Acesso em: 24 abr. 2017.

RIANI, J. L. R. *Determinantes do resultado educacional no Brasil: família, perfil escolar dos municípios e dividendo demográfico numa abordagem hierárquica e espacial*. 2005. 218f. Tese (Doutorado em Demografia) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.

RIANI, J. L. R.; SILVA, V. C. da; SOARES, T. M. Repetir ou progredir? Uma análise da repetência nas escolas públicas de Minas Gerais. *Educação e Pesquisa*, São Paulo, v. 38, n. 3, p. 623-636, jul./set. 2012.

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.

SHIRASU, M. R.; ARRAES, R. A. Determinantes da evasão e repetência escolar. In: ENCONTRO NACIONAL DE ECONOMIA, 43., 2015, Florianópolis. *Anais...* Florianópolis: Anpec, 2015.

SILVA, N. V.; HASENBALG, C. Recursos familiares e transições educacionais. *Cadernos de Saúde Pública*, Rio de Janeiro, v. 18, suplemento, p. 67-76, 2002.

SOARES, S. Os fatores que determinam o sucesso educacional. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 32, n. 3, p. 385-394, dez. 2002.

VASCONCELLOS, L. A relação entre frequência escolar e renda familiar no Brasil – 1981-1999. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 35, n. 2, p. 267-296, ago. 2005.

Recebido em: 29 AGOSTO 2017

Aprovado para publicação em: 31 AGOSTO 2018



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.