

# ESTUDO DE FATORES ASSOCIADOS ATRAVÉS DE REGRESSÃO QUANTÍLICA HIERÁRQUICA

**PEDRO ALBERTO BARBETTA<sup>I</sup>**  
**DALTON FRANCISCO DE ANDRADE<sup>II</sup>**  
**HÉLITON RIBEIRO TAVARES<sup>III</sup>**

## RESUMO

*Este trabalho apresenta uma abordagem não usual em estudo de fatores associados ao desempenho escolar, utilizando regressões quantílicas hierárquicas. Enquanto na abordagem tradicional tem-se por objetivo identificar fatores importantes com base nos indivíduos da faixa intermediária de proficiência, nesta aqui apresentada procura-se detectar efeitos nos quantis da distribuição de proficiência, permitindo identificar como determinado fator influencia estudantes de baixa, intermediária e alta proficiências. São realizadas a descrição da metodologia adotada e uma aplicação com os dados das provas de Língua Portuguesa e Matemática do 8º ano do ensino fundamental do estado do Pará, em 2016.*

<sup>I</sup> Universidade Federal de Santa Catarina (UFSC), Florianópolis-SC, Brasil; <http://orcid.org/0000-0002-5359-0134>; [pedro.barbetta@ufsc.br](mailto:pedro.barbetta@ufsc.br)

<sup>II</sup> Universidade Federal de Santa Catarina (UFSC), Florianópolis-SC, Brasil; <http://orcid.org/0000-0002-4403-980X>; [dalton.andrade@ufsc.br](mailto:dalton.andrade@ufsc.br)

<sup>III</sup> Universidade Federal do Pará (UFPA), Belém-PA, Brasil; <http://orcid.org/0000-0003-1725-3664>; [heliton@ufpa.br](mailto:heliton@ufpa.br)

**PALAVRAS-CHAVE** AVALIAÇÃO DA EDUCAÇÃO • FATORES ASSOCIADOS • REGRESSÃO QUANTÍLICA HIERÁRQUICA • RENDIMENTO ESCOLAR.

# ESTUDIO DE FACTORES ASOCIADOS A TRAVÉS DE REGRESIÓN CUANTÍLICA JERÁRQUICA

## RESUMEN

*Este trabajo presenta un abordaje no usual en el estudio de factores asociados al desempeño escolar, haciendo uso de regresiones cuantílicas jerárquicas. Mientras que en el enfoque tradicional se tiene el objetivo de identificar factores importantes en base a los individuos del rango intermedio de competencia, en este abordaje se pretende detectar efectos en los cuantiles de la distribución de competencia, lo que permite identificar cómo determinado factor influye en los estudiantes de competencia baja, intermedia y alta. Se describe la metodología adoptada y una aplicación con los datos de las pruebas de Lengua Portuguesa y Matemáticas del 8º año de la Educación Básica del estado de Pará, en el 2016.*

**PALABRAS CLAVE** EVALUACIÓN DE LA EDUCACIÓN • FACTORES ASOCIADOS • REGRESIÓN CUANTÍLICA JERÁRQUICA • RENDIMIENTO ESCOLAR.

# A STUDY ON ASSOCIATED FACTORS THROUGH HIERARCHICAL QUANTILE REGRESSION

## ABSTRACT

*In this paper, we present an unusual approach to factors associated with academic achievement since we use hierarchical quantile regressions. While the traditional approach aims to identify important factors based on individuals with an intermediate performance, our approach aims to detect the effects on performance distribution quantiles, thus allowing to identify how a certain factor can affect poor-, intermediate-, and high-performing pupils. We describe the methodology and use it with data from Portuguese and Mathematics tests of the 8th grade of primary education in the Brazilian state of Pará, in 2016.*

**KEYWORDS** EDUCATIONAL ASSESSMENT • ASSOCIATED FACTORS • HIERARCHICAL QUANTILE REGRESSION • ACADEMIC ACHIEVEMENT.

## **INTRODUÇÃO**

O conhecimento dos principais fatores que afetam o desempenho escolar auxilia os gestores e os professores a orientarem suas ações de forma mais objetiva e pragmática, melhorando o desempenho escolar dos estudantes e diminuindo as diferenças entre eles. Muitos fatores não dependem dos professores ou dos gestores, como o nível socioeconômico familiar, mas outros estão intimamente ligados à gestão e ao trabalho em sala de aula, como passar, cobrar e corrigir os deveres escolares dos estudantes. Identificar os fatores mais influentes no desempenho escolar e a forma que eles interagem com diferentes níveis de desempenho tem sido alvo de várias pesquisas científicas, sendo que a base das informações vem com as avaliações de larga escala.

Ao longo das últimas décadas no Brasil, nota-se que a avaliação em larga escala vem se consolidando como um instrumento das políticas públicas em educação, em seus diferentes âmbitos de gestão, seja ele em nível federal, estadual ou municipal. No entanto, ainda se observam inúmeros desafios quanto ao uso efetivo dessas avaliações na

formulação, reformulação e no monitoramento de ações, projetos e programas educacionais, e, notadamente, no que diz respeito ao trabalho no interior das unidades escolares. (GIMENES, 2015, p. 254)

Nos estudos de fatores associados ao desempenho escolar, seja nos níveis fundamental, médio ou superior, a metodologia de modelos hierárquicos se destaca, já que existem fatores associados ao estudante e fatores relacionados à escola onde ele estuda, ou seja, há um nível de hierarquia primário – os estudantes – e um nível agregado – as escolas (ALVES; SOARES, 2008; ANDRADE; SOARES, 2008; LAROS; MARCIANO; ANDRADE, 2010; NASCIMENTO, 2012; MORICONI; NASCIMENTO, 2014; PONTES; SOARES, 2016). Como o desempenho de um estudante numa avaliação de larga escala é, geralmente, medido por uma escala contínua, a maioria dos estudos usa modelos hierárquicos lineares.

Os modelos lineares hierárquicos podem ser considerados uma extensão da regressão linear múltipla para acomodar uma estrutura hierárquica, como é o caso dos estudantes que estão agrupados em escolas. Nesses modelos, procura-se avaliar a parcela do desempenho do estudante devida a cada fator associado ao estudante (nível socioeconômico familiar, percepção do estudante quanto ao clima escolar, necessidade de o estudante fazer trabalhos domésticos, etc.) e cada fator associado à escola (infraestrutura, localização, dependência administrativa, etc.).

Um dos principais objetivos do estudo de fatores associados está na avaliação da influência da escola no desempenho cognitivo de seus alunos, conhecido como *efeito escola* (ALVES; SOARES, 2013; ANDRADE; SOARES, 2008). Esse efeito é obtido com a retirada da influência dos fatores extraescolares, como o nível socioeconômico e ambiente da região da escola, resultando na parcela relativa ao ambiente escolar propriamente dito, o qual pode ser aprimorado por uma boa gestão escolar.

A questão que se coloca é que tradicionalmente os modelos relacionam o valor médio do desempenho dos estudantes com os diversos fatores extra e intraescolares. Mas

muitos desses fatores podem afetar diferentemente estudantes de baixo e de alto desempenho. Nas políticas educacionais que buscam melhorar o aprendizado de crianças com maiores dificuldades na escola, há interesse em encontrar os fatores intraescolares com maior efeito em crianças com baixo desempenho.

Este trabalho propõe avaliar o efeito dos diferentes fatores associados não só na média de proficiência dos estudantes, mas também em diversos quantis (ou percentis) da distribuição das proficiências, permitindo verificar se determinado fator tem maior efeito em estudantes (ou em escolas) de alto desempenho ou naqueles (ou em escolas) de baixo desempenho. Promover um fator de ambiente escolar com efeito positivo e mais forte em estudantes (ou em escolas) de baixo desempenho é colaborar com a equidade do aprendizado.

A metodologia proposta é aplicada no Sistema Paraense de Avaliação Educacional (SisPAE), usando as provas do 8º ano do ensino fundamental (EF) e medidas e indicadores obtidos nos questionários de contexto.

### **QUANTIS, REGRESSÃO QUANTÍLICA E ESTRUTURA HIERÁRQUICA**

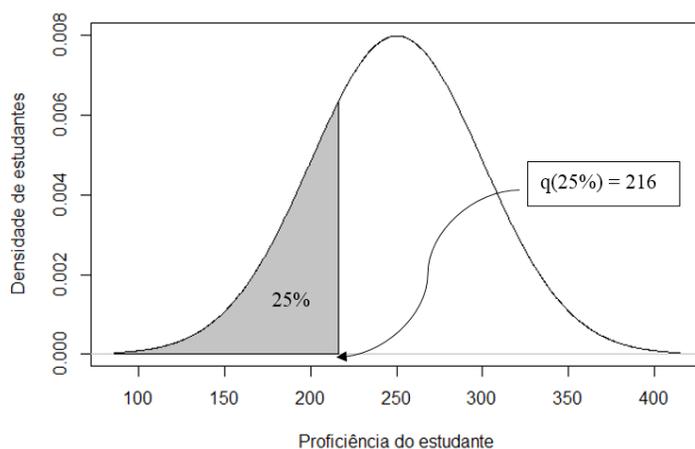
A metodologia tradicional para análise de fatores associados ao desempenho escolar envolve a análise de regressão, que tem como foco identificar a influência de cada fator na média da distribuição do desempenho. Usualmente os parâmetros da equação de regressão são obtidos por mínimos quadrados ordinários (MQO).<sup>1</sup>

Muitas vezes existe interesse em verificar se determinado fator afeta mais fortemente os estudantes de baixo ou de alto desempenho. Uma forma de fazer a análise, nesse caso, é por meio da chamada regressão quantílica (KOENKER, 2005). Por esse procedimento, a análise pode ser em termos de quantis (ou percentis) predefinidos, como a mediana, primeiro quartil e terceiro quartil da distribuição de desempenho dos estudantes. Se houver apenas uma variável independente, a regressão quantílica, basicamente, estima várias retas para diferentes quantis de desempenho desejados.

<sup>1</sup> Para detalhes, ver, por exemplo, Neter *et al.* (1996).

O quantil mais conhecido é a mediana,  $q(50\%)$ , que separa os 50% menores valores dos 50% maiores valores. Numa distribuição de proficiências de avaliados, o  $k$ -ésimo quantil, aqui denominado de  $q(k)$ , pode ser entendido como o valor que delimita os  $k\%$  menores valores dos  $(100 - k)\%$  maiores valores. A Figura 1 ilustra a posição do quantil 25%, denotado por  $q(25\%)$ , que corresponde ao valor na escala que separa 25% dos estudantes com menor proficiência, numa distribuição hipotética. O quantil  $q(25\%)$  também é conhecido como primeiro quartil ou quartil inferior.

**FIGURA 1 - Ilustração do quantil 25%: ponto que separa 25% dos estudantes com menor proficiência dos 75% com maior proficiência**

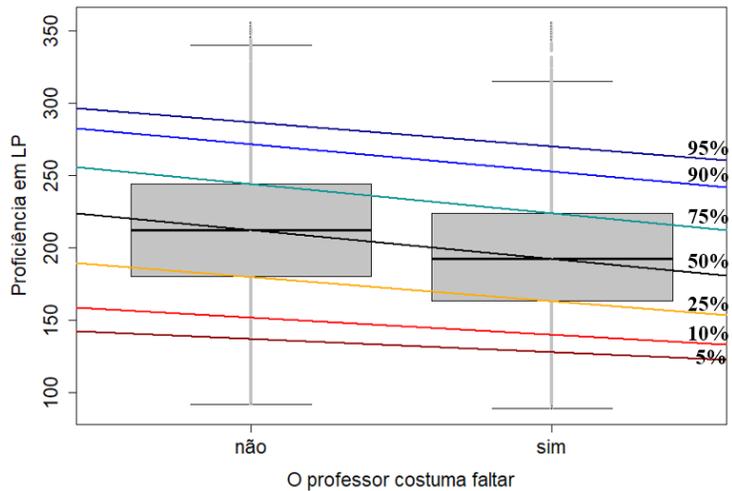


Fonte: Elaboração dos autores, 2018.

Uma vantagem de usar a regressão quantílica para a mediana, em vez da regressão de mínimos quadrados ordinários para a média, é que o seu resultado será mais robusto, em resposta aos *outliers*. Chen *et al.* (2014) apontam que a regressão quantílica é mais útil do que a regressão da média quando o termo aleatório não tem distribuição normal. Esse tipo de regressão pode ser visto como uma analogia natural à prática de usar diferentes medidas de tendência central, dispersão e assimetria para obter uma análise mais abrangente e mais robusta (KOENKER, 2005). Cabe ainda ressaltar que na regressão quantílica qualquer quantil pode ser analisado.

A Figura 2 ilustra vários quantis da distribuição de proficiência de estudantes na comparação de dois grupos, definidos em termos de professores assíduos e professores que faltam muito, conforme respostas do questionário aplicado aos alunos. Nessa figura, os quantis são apresentados por diagramas em caixas – o que constitui uma representação usual de medianas, quartis e extremos – e por regressões quantílicas com a variável independente indicadora do atributo em questão (0 = professores assíduos; 1 = professores que faltam muito).

**FIGURA 2 - Representação gráfica da distribuição das proficiências em Língua Portuguesa de estudantes do 8º ano do ensino fundamental que foram considerados segundo a assiduidade (ou não) de seus professores, vista em termos de quantis**



Fonte: Elaboração dos autores, 2018.

Com regressões quantílicas, como ilustrado na Figura 2, além de se avaliarem possíveis diferenças na posição central da distribuição, é possível também avaliar se o efeito do fator de interesse é maior nos estudantes de baixo desempenho ou naqueles de alto desempenho. No exemplo em questão, percebe-se que o desempenho (mediano) esperado de estudantes cujo professor costuma faltar é inferior ao daqueles cujo professor não costuma faltar. No entanto, parece que o efeito do professor faltar é aproximadamente

igual nos quantis analisados (retas aproximadamente paralelas). Situações em que as retas nos quantis inferiores são mais inclinadas (têm coeficientes angulares com magnitudes maiores) do que em quantis superiores indicam que o fator é mais influente nos estudantes de baixo desempenho. Martins e Pereira (2004), Davino e Vistocco (2008), Rangvid (2008) e Almeida (2014) apresentam aplicações de regressão quantílica na área da educação.

Outra característica, bastante presente na análise de dados educacionais, é a estrutura hierárquica dos mesmos, ou seja, os estudantes são subunidades de escolas, que podem ser subunidades de diretorias, regionais ou outros agrupamentos. Com isso temos agrupamentos de estudantes dentro de escolas, gerando possivelmente medidas correlacionadas dentro de cada grupo. Temos assim uma fonte de variação interna às escolas e outra que representa a variação entre as escolas. A análise pode ser afetada pelas correlações intra-grupos, de tal forma que os modelos hierárquicos retratam bem essa estrutura de dados.

Enfim, com o objetivo de verificar se um fator afeta mais os estudantes com desempenho escolar alto ou aqueles com desempenho escolar baixo, este trabalho foca nos quantis da distribuição de proficiência, considerando a estrutura hierárquica dos dados. As regressões múltiplas com modelagem hierárquica permitem isolar o efeito de cada fator e separar a variação entre escolas e a variação entre estudantes de uma mesma escola.

Nos modelos tradicionais, os efeitos dos fatores são avaliados em termos de diferenças nos valores esperados. O presente trabalho inova no sentido de estudar as diferenças de efeitos em termos de quantis, permitindo avaliar os efeitos em diferentes níveis do desempenho acadêmico.

## **METODOLOGIA**

Em termos teóricos, o  $k$ -ésimo quantil,  $q(k)$ , é dado pela solução da equação  $P\{Y \leq q(k)\} = k$ , sendo  $Y$  uma variável aleatória contínua. Koenker (2005) e Koenker e Bassett (1978) apresentam uma maneira de se obter a mediana,  $q(50\%)$ , pela técnica

dos mínimos quadrados. Com base em um conjunto de valores  $(y_1, y_2, \dots, y_n)$ , obtém-se o valor de  $\theta$  que minimiza a soma:

$$S(\theta) = \sum_i |y_i - \theta| \quad (1)$$

Para os demais quantis,  $q(k)$ , estende-se a relação considerando pesos diferentes para os resíduos positivos e negativos:

$$S(\theta) = \sum_i \rho_k(y_i - \theta), \text{ sendo } \rho_k(r_i) = \begin{cases} k |r_i| & r_i > 0 \\ (1-k) |r_i| & r_i \leq 0 \end{cases} \quad (2)$$

O quantil  $q(k)$  é o valor de  $\theta$  que minimiza a relação (2). Esse método de mínimos erros absolutos é equivalente ao de máxima verossimilhança quando se supõe que os erros têm distribuição de Laplace, conforme descrito em Geraci e Bottai (2014). Segundo Yu e Zhang (2006), a distribuição de Laplace é definida por três parâmetros: um de locação, um de dispersão e um de assimetria, o que permite realizar a análise para situações em que a distribuição da variável em estudo é simétrica ou assimétrica.

Com base em Geraci e Bottai (2014), descreve-se o modelo de regressão quantílica hierárquica, conforme será usado neste trabalho:

$$Y_{ij} = \beta_{0j}^{(k)} + \beta_1^{(k)} X_{1ij} + \beta_2^{(k)} X_{2ij} + \dots + \beta_l^{(k)} X_{lij} + e_{ij}^{(k)} \text{ (nível 1)}$$

com

$$\beta_{0j}^{(k)} = \gamma_0^{(k)} + \gamma_1^{(k)} W_{1j} + \gamma_2^{(k)} W_{2j} + \dots + \gamma_M^{(k)} W_{Mj} + u_j^{(k)} \text{ (nível 2) } (3)$$

sendo:

$Y_{ij}$  a variável dependente (proficiência) do estudante  $i$  da escola  $j$ ;

$\beta_{0j}^{(k)}$  o  $k$ -ésimo quantil da escola  $j$ , condicionado às demais variáveis do modelo iguais a zero;

$\beta_l^{(k)}$  o efeito da  $l$ -ésima variável de estudante no  $k$ -ésimo quantil ( $l = 1, 2, \dots, L$ );

$\gamma_0^{(k)}$  o  $k$ -ésimo quantil médio das escolas, condicionado às demais variáveis do modelo iguais a zero;

$\gamma_m^{(k)}$  o efeito da  $m$ -ésima variável de escola no  $k$ -ésimo quantil ( $m = 1, 2, \dots, M$ );

$e_{ij}^{(k)}$  o erro aleatório associado ao  $i$ -ésimo estudante da escola  $j$ , supostamente com distribuição de Laplace com parâmetro de locação igual a zero, de dispersão igual a  $\sigma$  e de assimetria igual a  $k$ ;

$u_j^{(k)}$  o termo aleatório associado à escola  $j$ , supostamente com distribuição normal de média zero e variância  $\sigma_u^2$ .

Neste trabalho está se controlando a heterogeneidade das escolas pelo efeito aleatório em  $\beta_{0j}^{(k)}$ , ou seja, cada escola pode ter o  $k$ -ésimo quantil diferente. Num contexto geral, poderia haver efeito aleatório em alguns dos efeitos das variáveis de estudantes, ou seja, nos  $\beta_l^{(k)}$  ( $l = 1, 2, \dots, L$ ). Por exemplo, poder-se-ia supor que o efeito da reprovação seja diferente, dependendo da escola, já que algumas escolas podem ter tratamentos diferentes para os estudantes que não lograram aprovação.

Os parâmetros do modelo podem ser estimados pelo método da máxima verossimilhança. Para isso será usado o pacote *lqmm* (GERACI, 2016) do *software* livre R (R FOUNDATION FOR STATISTICAL COMPUTING, 2017).<sup>2</sup>

A literatura sobre o assunto não é muito extensa, mas há outras propostas para introduzir estrutura hierárquica na regressão quantílica, como em Chen *et al.* (2014) e Hassan (2014). Vale destaque a abordagem de Galarza, Lachos e Bandyopadhyay (2017), que propõem melhorias no algoritmo para obtenção das estimativas por máxima verossimilhança em regressão quantílica hierárquica, oferecendo novo pacote no R, *qrLMM*, em estudo pelos autores do presente trabalho.

## DADOS E VARIÁVEIS DO ESTUDO

A metodologia proposta é aplicada no sistema de avaliação em larga escala do estado do Pará (SisPAE), iniciado em 2013, que avalia os estudantes nas áreas de Língua Portuguesa e Matemática. Na edição 2016 foram avaliados alunos do 4º e 8º anos do ensino fundamental, bem como as 3<sup>as</sup> séries do

<sup>2</sup> Mais referências sobre esse método podem ser obtidas em Geraci e Bottai (2007, 2014), Geraci (2014) e Bottai, Orsini e Geraci (2015).

ensino médio. A avaliação envolveu a participação de alunos, professores e diretores de 3.534 escolas, distribuídas nos 144 municípios paraenses. São 814 escolas estaduais, 2.714 municipais e seis casas familiares rurais. Da previsão inicial de 563.413 alunos matriculados, participaram efetivamente da avaliação 376.830 em Língua Portuguesa e 376.684 em Matemática.

A edição 2016 do SisPAE incluiu coleta de informações sobre as características dos alunos de todos os anos/séries avaliados, bem como dos professores, orientadores pedagógicos, especialistas em educação, diretores e escolas, com o objetivo de traçar o perfil dos respondentes e coletar dados para uma análise mais detalhada dos fatores associados ao desempenho escolar. Para tanto, foram aplicados questionários nos seguintes quantitativos: 270.424 alunos; 15.250 professores; 3.094 orientadores pedagógicos; 19.769 especialistas em educação; 3.509 diretores; e 3.509 escolas.

No estudo de fatores associados do SisPAE (PARÁ, 2017), são descritas várias medidas e indicadores construídos com base nos questionários contextuais, especialmente naqueles respondidos pelos estudantes. A escolha das medidas e indicadores baseou-se na literatura de fatores associados ao desempenho escolar, como em Alves e Soares (2008, 2013) e Laros, Marciano e Andrade (2010), bem como em relatórios técnicos de avaliações de larga escala efetuadas em outras unidades da federação. No presente artigo, foram selecionados os indicadores que apresentaram significância estatística, aqueles mais fortemente associados com a proficiência, conforme a análise de SisPAE (PARÁ, 2017). Como o objetivo central deste trabalho é mostrar uma nova abordagem de análise, optou-se por não usar número excessivo de fatores.

Em SisPAE 2016 (PARÁ, 2017) foram construídas várias medidas pela Teoria da Resposta ao Item (TRI) com modelo de escala gradual (ANDRADE; TAVARES; VALLE, 2000; AYALA, 2009), das quais três foram usadas neste artigo:

- Nível Socioeconômico da família do estudante (NSE);
- Nível de Relacionamento Escolar na percepção do estudante (NRE);
- Nível do Clima Escolar na percepção do estudante (NCE).

As medidas também foram agregadas, em termos de médias, para a(s) turma(s) do ano escolar do estudante, em sua escola. Neste artigo, tais medidas foram padronizadas com média zero e desvio-padrão um, relativo ao grupo de estudantes em estudo. No caso da medida agregada, a padronização foi em relação ao grupo de escolas pesquisadas.

Também foram consideradas variáveis indicadoras da presença de fatores associados ao desempenho escolar: sexo do estudante; indicador de reprovação; indicador de realização de dever de casa; dentre outros. A lista completa dos fatores aqui analisados encontra-se na primeira coluna da Tabela 1.

## **RESULTADOS E DISCUSSÃO**

Nesta seção são apresentados e discutidos os resultados de regressões quantílicas hierárquicas aplicadas nas provas de Língua Portuguesa e Matemática do 8º ano do ensino fundamental do estado do Pará, em 2016.

A obtenção das estimativas teve por base um conjunto de 1.808 escolas e mais de 66.000 estudantes, que correspondem às unidades observacionais com respostas nas variáveis em estudo.

Para melhor ilustrar as informações que podem ser extraídas com a abordagem proposta, será apresentada, inicialmente, uma análise pela regressão hierárquica usual, baseada no valor médio da proficiência. A Tabela 1 traz a estimativa do efeito esperado por cada fator incluído na análise. O método de estimação foi o de máxima verossimilhança restrita, usando o pacote lme4 (BATES et al., 2015).

**TABELA 1 - Estimativas dos efeitos esperados pelos fatores em análise, ajustando-se um modelo hierárquico aos dados do 8º ano do ensino fundamental do SisPAE 2016, por disciplina**

VARIÁVEIS / FATORES	LÍNGUA PORTUGUESA			MATEMÁTICA		
	ESTIMATIVAS	E.P.	ESTATÍSTICA T	ESTIMATIVAS	E.P.	ESTATÍSTICA T
Intercepto	200,6	0,8	260,6	215,9	0,6	367,4
<b>NÍVEL 1 - ESTUDANTE</b>						
Nível Socioeconômico da família do estudante (NSE)	3,6	0,2	21,2	3,3	0,1	22,3
Nível de Relacionamento Escolar estudante (NRE)	1,7	0,2	10,5	1,4	0,1	9,7
Nível do Clima Escolar estudante (NCE)	0,5	0,2	2,8	0,7	0,1	4,4
Sexo feminino	10,6	0,3	33,5	-5,0	0,3	-18,4
Estudante teve reprovação	-12,8	0,3	-39,1	-10,5	0,3	-37,6
Tempo gasto até a escola maior que 1/2 hora	-6,9	0,4	-16,5	-4,5	0,4	-12,7
Estudante aprende matéria com professor	5,2	0,5	10,8	5,9	0,4	14,4
Aprendizado depende mais do professor do que do estudante	-9,8	0,3	-28,7	-6,9	0,3	-23,7
Professor dá mais atenção aos estudantes bons	-6,2	0,4	-17,7	-4,2	0,3	-14,1
Professor se preocupa com o dever de casa	8,9	0,4	25,3	5,0	0,3	16,6
Professor costuma faltar	-5,4	0,3	-15,9	-3,7	0,3	-12,7
Estudante lê frequentemente	3,7	0,4	9,7	-	-	-
Pais incentivam a ler	4,3	0,4	10,4	-	-	-
<b>NÍVEL 2 - ESCOLA</b>						
Nível Socioeconômico do ano/escola	4,4	0,3	15,6	2,7	0,3	10,9
Nível de Relacionamento Escolar da série/escola	1,2	0,3	4,0	1,2	0,3	4,6
Nível do Clima Escolar da série/escola	2,0	0,3	6,6	2,1	0,3	8,0

Fonte: Elaboração dos autores, 2018.

Todas as estimativas apresentadas na Tabela 1 são significantes estatisticamente ( $p < 0,05$ ).<sup>3</sup> Analisando as estimativas dos efeitos, chega-se à conclusão de que, na população estudada, os níveis de clima e de relacionamento escolar (NCE e NRE, respectivamente) têm efeitos positivos no desempenho escolar. Em números, a cada desvio-padrão a mais em que os estudantes percebem clima escolar da escola, a proficiência média em Língua Portuguesa tende a ser duas unidades superior.

Em geral, as meninas foram melhores em Língua Portuguesa e os meninos em Matemática: estimativa do efeito nas médias de meninas em relação aos meninos é de 10,6 em Língua Portuguesa e -5,0 em Matemática. O grupo de es-

<sup>3</sup> Todos os valores da estatística *t* apresentados na tabela têm magnitudes maiores do que dois. Como as quantidades de estudantes e de escolas são bastante grandes, o valor *p* pode ser obtido por áreas nas caudas da distribuição normal. A área agregada acima de 2 e abaixo de -2 é inferior a 5%, ou seja,  $p < 0,05$ .

tudantes que tiveram reprovação obteve proficiência média bem inferior aos outros: estimativas de -12,8 para Língua Portuguesa e -10,5 para Matemática, ou seja, a retenção do estudante devido ao seu mal desempenho (ou outros motivos) não tende a fazer com que este chegue ao nível dos demais estudantes nos anos seguintes.

Os resultados citados no parágrafo anterior são compatíveis com outros estudos de fatores associados, como em Laros, Marciano e Andrade (2010) e Alves e Soares (2008). O que se coloca é que esses efeitos são observados no nível médio de proficiência, deixando de responder questões tais como: será que o clima escolar pode influenciar mais os estudantes de baixa proficiência? Será que a diferença entre o sexo feminino e o masculino acontece com maior ênfase em estudantes de baixa ou de alta proficiência? E o efeito do professor se preocupar com os deveres de casa dos estudantes é maior entre aqueles de baixo ou de alto desempenho escolar? De modo geral, o que se tenta responder neste artigo é se determinado efeito é maior em estudantes de alto desempenho ou de baixo desempenho. Para tanto, a análise que segue baseia-se nos quantis da distribuição das proficiências.

A Tabela 2 apresenta vários quantis dos resultados das provas do SisPAE 2016, na escala Saeb. Considerando a prova de Língua Portuguesa, ao analisar o quantil 10%, igual a 147,4, diz-se que 10% dos estudantes têm proficiência em Língua Portuguesa igual ou inferior a 147,4 pontos. Em termos do quantil 90%, pode-se afirmar que 90% dos estudantes têm proficiência em Língua Portuguesa igual ou inferior a 256,1 pontos, ou, equivalentemente, 10% dos estudantes têm proficiência superior a 256,1 pontos. Ou seja, enquanto  $q(10\%)$  separa os 10% de menor proficiência,  $q(90\%)$  separa os 10% de maior proficiência.

**TABELA 2 - Proficiência dos alunos do 8º ano do ensino fundamental do SisPAE 2016, por disciplina, considerando-se alguns quantis de interesse**

DISCIPLINA	QUANTIL				
	10%	25%	50%	75%	90%
Língua Portuguesa	147,4	171,8	201,2	231,1	256,1
Matemática	162,7	181,3	203,9	228,4	252,2

Fonte: Elaboração dos autores, 2018.

A discussão nesta seção envolve o quanto diferenças de valores nas variáveis explicativas – ou da presença de fatores que costumam afetar o desempenho – alteram os valores dos quantis e, mais ainda, se o efeito é maior (ou menor) nos quantis que separam os estudantes de desempenho mais baixo ou nos quantis que separam aqueles de desempenho mais alto.

A Tabela 3 traz as estimativas dos efeitos dos fatores em análise sobre alguns quantis da distribuição de proficiência em Língua Portuguesa. No geral, serão apresentadas apenas as estimativas de máxima verossimilhança que registraram significância estatística ao nível de 0,05. Tabelas com erros-padrão e testes de significância desses quantis encontram-se no Apêndice (tabelas A1 a A5).

**TABELA 3 - Coeficientes dos modelos de regressão quantílica hierárquica ajustados aos resultados da prova de Língua Portuguesa do 8º ano do ensino fundamental do SisPAE 2016**

VARIÁVEIS/FATORES	QUANTIL				
	10%	25%	50%	75%	90%
Intercepto	145,8	168,2	199,4	227,2	259,7
<b>NÍVEL 1 - ESTUDANTE</b>					
Nível Socioeconômico da família do estudante (NSE)	2,5	3,1	3,7	5,1	4,6
Nível de Relacionamento Escolar estudante (NRE)	1,9	2,0	2,0	1,6	1,3
Nível do Clima Escolar estudante (NCE)	-	-	0,6	1,1	-
Sexo feminino	11,8	13,0	11,7	10,4	6,8
Estudante teve reprovação	-9,3	-11,2	-13,0	-12,9	-15,1
Tempo gasto até a escola maior que 1/2 hora	-6,7	-7,7	-7,7	-5,4	-5,4
Estudante aprende matéria com professor	8,2	7,1	5,4	2,9	-
Aprendizado depende mais do professor do que do estudante	-6,1	-8,1	-10,4	-11,7	-12,4
Professor dá mais atenção aos estudantes bons	-5,2	-6,1	-6,4	-6,2	-7,4
Professor se preocupa com o dever de casa	6,3	8,8	9,5	11,1	10,3
Professor costuma faltar	-5,3	-5,5	-5,5	-4,3	-4,9
Estudante lê frequentemente	3,0	3,9	3,2	4,9	4,2
Pais incentivam a ler	3,3	4,0	5,3	4,7	3,8
<b>NÍVEL 2 - ESCOLA</b>					
Nível Socioeconômico do ano/escola	2,4	3,2	4,1	4,7	6,6
Nível de Relacionamento Escolar da série/escola	1,1	1,7	-	1,6	-
Nível do Clima Escolar da série/escola	1,7	2,5	2,3	1,2	1,8

Fonte: Elaboração dos autores, 2018.

O efeito do nível socioeconômico familiar (NSE) no desempenho escolar é deveras conhecido. Pela Tabela 3, verifica-se que o aumento de um desvio-padrão do NSE da família do estudante leva a um aumento de 3,7 pontos na proficiência mediana, estimativa parecida com o que foi obtido com o modelo para a média (Tabela 1), o que é razoável, já que ambas as medidas são de tendência central da distribuição de proficiência. Porém, com as regressões quantílicas, observa-se que o efeito do NSE é mais forte em estudantes com proficiências mais altas: enquanto no quantil 10% (ponto que separa 10% dos estudantes com menor proficiência) o efeito estimado de um desvio-padrão no NSE foi de 2,5 pontos a

mais na prova; no quantil 90%, que separa 10% dos estudantes com maior proficiência, o efeito estimado foi de 4,6 pontos. Analisando os valores do NSE ano/escola, verifica-se que a relação é similar ao que ocorre com o NSE da família do estudante. A Tabela 4 mostra os quantis da distribuição das proficiências de Matemática. As relações são bastante parecidas com Língua Portuguesa para a maioria das variáveis/fatores analisados.

**TABELA 4 - Coeficientes dos modelos de regressão quantílica hierárquica ajustados aos resultados da prova de Matemática do 8º ano do ensino fundamental do SisPAE 2016**

VARIÁVEIS/FATORES	QUANTIL				
	10%	25%	50%	75%	90%
Intercepto	168,9	187,9	213,7	240,7	266,1
<b>NÍVEL 1 - ESTUDANTE</b>					
Nível Socioeconômico da família do estudante (NSE)	1,9	2,6	3,4	4,3	4,2
Nível de Relacionamento Escolar estudante (NRE)	1,4	1,4	1,5	1,4	1,1
Nível do Clima Escolar estudante (NCE)	-	-	0,6	1,0	1,6
Sexo feminino	-0,9	-1,9	-4,4	-7,2	-9,6
Estudante teve reprovação	-6,2	-8,0	-10,3	-12,2	-14,3
Tempo gasto até a escola maior que 1/2 hora	-3,7	-4,5	-4,9	-5,0	-5,3
Estudante aprende matéria com professor	4,2	5,3	6,0	5,9	5,9
Aprendizado depende mais do professor do que do estudante	-3,5	-4,9	-6,9	-8,6	-9,5
Professor dá mais atenção aos estudantes bons	-2,8	-3,7	-4,5	-5,0	-5,4
Professor se preocupa com o dever de casa	3,1	4,5	5,3	5,5	5,6
Professor costuma faltar	-2,2	-3,1	-3,4	-4,1	-3,9
<b>NÍVEL 2 - ESCOLA</b>					
Nível Socioeconômico do ano/escola	1,9	2,5	2,9	2,7	4,4
Nível de Relacionamento Escolar da série/escola	0,8	1,3	1,2	1,4	-
Nível do Clima Escolar da série/escola	1,6	1,8	2,2	2,7	3,5

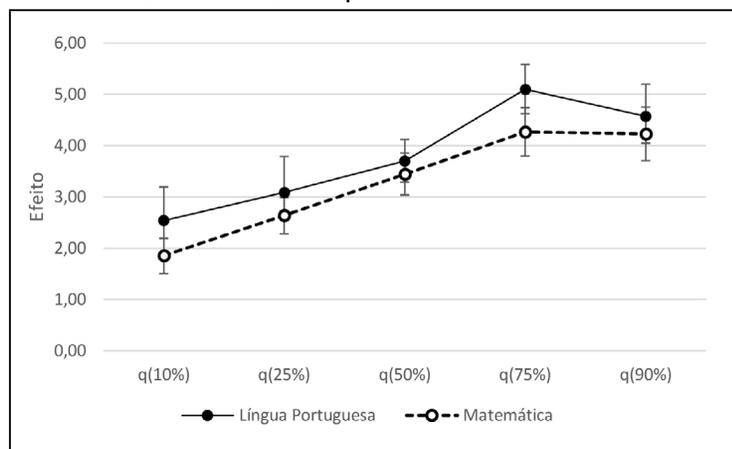
Fonte: Elaboração dos autores, 2018.

Nota: Os fatores relacionados à leitura (incentivo dos pais e se o próprio estudante costuma ler) foram excluídos por não apresentarem efeitos significantes e nem terem muito sentido para a prova de Matemática.

As figuras 3 a 7 reproduzem as estimativas dos efeitos em cada quantil para os fatores em que foram observadas diferenças importantes, Língua Portuguesa e Matemática, incluindo intervalos de 95% de confiança.

A Figura 3 mostra os efeitos do nível socioeconômico familiar. Observa-se que, em ambas as áreas, os intervalos nos quantis 10% e 90% são disjuntos, indicando diferença significativa; o mesmo ocorre entre os quartis inferior,  $q(25\%)$ , e superior,  $q(75\%)$ . Ou seja, o efeito do NSE é significativamente mais forte em estudantes com proficiências mais altas.

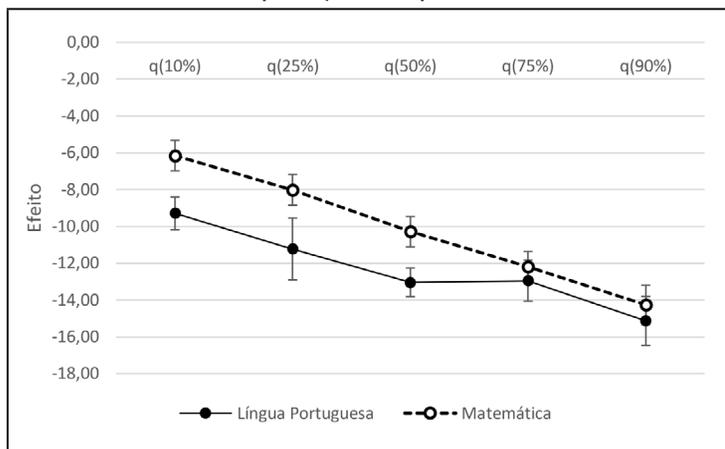
FIGURA 3 - Efeito do NSE nos quantis analisados



Fonte: Elaboração dos autores, 2018.

Um fator bastante estudado nas avaliações de larga escala é o da reprovação. Em ambas as provas se verifica efeito negativo do estudante que já teve reprovação, com redução de mais de 10 pontos na mediana, quando comparado ao grupo que não teve reprovação. Percebe-se, ainda, que esse efeito é mais forte dentre os estudantes com maior desempenho. No quantil 90%, que separa 10% dos estudantes com maior desempenho, a diferença entre os dois grupos foi de 15,1 pontos em Língua Portuguesa e 14,3 pontos em Matemática (Figura 4). Cabe ressaltar que há bem menos casos de estudantes de alto desempenho com reprovação, mas se deve considerar, também, a influência de algum outro fator não controlado, tal como doença do estudante ou em sua família em anos anteriores, levando à retenção.

FIGURA 4 - Efeito da reprovação nos quantis analisados

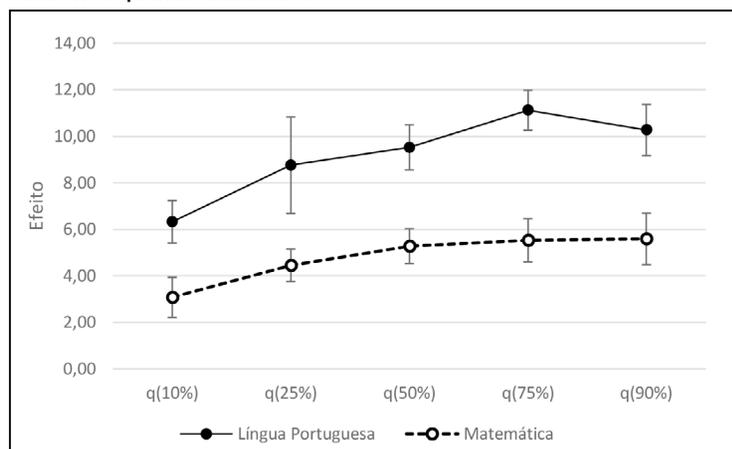


Fonte: Elaboração dos autores, 2018.

São identificados efeitos negativos associados ao “professor que costuma faltar” e ao “professor que dá mais atenção aos estudantes bons” (tabelas 3 e 4), na percepção do estudante respondente. Não há uma tendência clara de a magnitude do efeito aumentar ou diminuir à medida que se caminha para quantis maiores, ou seja, os efeitos desses indicadores são igualmente ruins entre os estudantes de melhor e de pior desempenho.

O efeito positivo do “professor que se preocupa com o dever de casa” é ligeiramente maior nos estudantes com melhor proficiência. Isso sugere que os estudantes de melhor desempenho aproveitam melhor os deveres de casa e as correções feitas pelo professor (Figura 5).

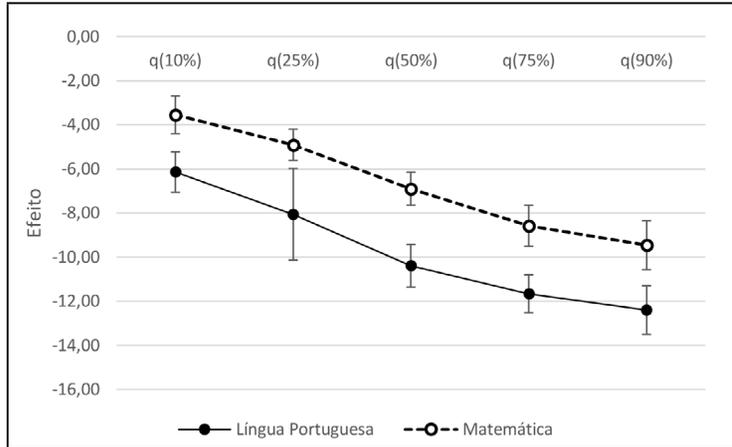
**FIGURA 5 - Efeito de o “professor se preocupar com o dever de casa” nos quantis analisados**



Fonte: Elaboração dos autores, 2018.

Um aspecto interessante é o forte efeito negativo dos “estudantes que consideram que o aprendizado depende mais do professor do que do estudante”. Com as regressões quantílicas pode-se perceber que esse efeito é mais forte dentre os estudantes de desempenho maior. Em Língua Portuguesa, para estudantes que estão no quantil 90%, a diferença é em torno de 12 pontos. Já para aqueles que estão no quantil 10%, o efeito negativo é em torno de 6 pontos. Talvez esse comportamento seja característico de alunos de menor desempenho, cuja homogeneidade dos grupos (concordantes e discordantes da pergunta) leve a uma menor diferença de desempenho, ocorrendo o oposto para os estudantes de maior desempenho (Figura 6).

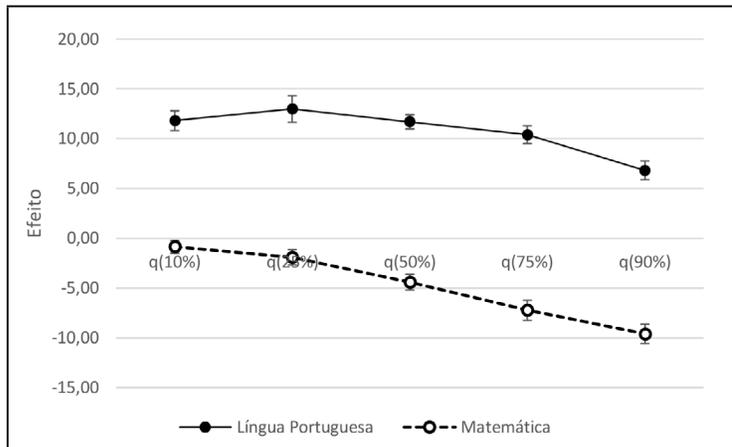
**FIGURA 6 - Efeito da concordância pelos estudantes de que “o aprendizado depende mais do professor do que dele próprio” nos quantis analisados**



Fonte: Elaboração dos autores, 2018.

Um aspecto interessante, verificado em vários estudos de avaliação de larga escala, é que em Matemática os meninos têm desempenho melhor, enquanto em Língua Portuguesa são as meninas que mostram maior desempenho. A Figura 7 apresenta os efeitos do fator sexo em diferentes quantis, tanto em Língua Portuguesa quanto em Matemática.

**FIGURA 7 - Efeito do sexo feminino nos quantis analisados**



Fonte: Elaboração dos autores, 2018.

A novidade observada neste trabalho é que em Matemática a diferença entre os sexos é maior entre os estudantes de mais alto desempenho, com os meninos tendo desempenho bem superior ao das meninas (quase 10 pontos no quantil 90%). Por outro lado, em Língua Portuguesa, em que as meninas sobressaem, a diferença é maior entre os estudantes de baixo desempenho (mais de 10 pontos nos quantis inferiores). Pensando conjuntamente em Língua Portuguesa e Matemática, a análise sugere que os meninos estão em posição mais extrema do que as meninas, já que em Língua Portuguesa, disciplina em que eles usualmente têm desempenho pior, a diferença é maior dentre aqueles de pior desempenho; em Matemática, em que eles normalmente têm desempenho superior, a diferença acentua entre os de maior desempenho.

## **CONCLUSÃO**

Há poucos trabalhos sobre estudos de fatores associados ao desempenho escolar usando regressão quantílica, e em especial regressão quantílica hierárquica. De certa forma isso é justificável, porque regressão quantílica hierárquica ainda não é um método suficientemente conhecido e consagrado. Outro fator que justifica os poucos trabalhos com esse método é que ele não está implementado nos principais *softwares* comerciais da área de estatística. Contudo, as análises aqui realizadas mostraram boas evidências de fatores que exercem influência mais expressiva na comparação de estudantes de alto e baixo desempenhos, resultando num importante subsídio para a gestão escolar.

Neste trabalho foi mostrado, por exemplo, que o nível socioeconômico em que o estudante está inserido provoca maiores diferenças em estudantes de alto desempenho, o que também ocorre com o efeito positivo de o professor se preocupar com os deveres de casa. Estudantes com desempenho baixo parecem mais dependentes do professor, pois a concordância de que o aprendizado depende mais do professor do que do aluno tem alto efeito negativo nos vários quantis da distribuição. Com a abordagem da regressão quantílica, verificou-se que esse efeito negativo é mais forte

nos grupos de desempenho mais alto, ou seja, estudantes com bom desempenho que concordam com essa afirmação tendem a ficar bem abaixo daqueles com bom desempenho que discordam.

Os resultados devem encorajar novos estudos em outras populações que possam corroborar várias relações discutidas neste trabalho. Mais ainda, há políticas públicas realizadas com o objetivo de atenuar as diferenças de desempenho entre os estudantes, tentando melhorar o desempenho especialmente daqueles com maior dificuldade de aprendizado. A regressão quantílica hierárquica torna-se um bom instrumento para avaliar de forma objetiva o efeito dessas políticas, porque mede o efeito dos fatores de interesse em diferentes níveis do desempenho educacional, não só em seu valor esperado.

## **AGRADECIMENTOS**

À Fundação para o Vestibular da Unesp (Vunesp), pela cessão dos dados e pelo incentivo no desenvolvimento de pesquisas em avaliação educacional.

## **REFERÊNCIAS**

ALMEIDA, A. T. C. Determinante dos piores e melhores resultados educacionais dos alunos da rede pública de ensino fundamental no Brasil. *Planejamento de Políticas Públicas*, Brasília, DF, n. 42, p. 147-187, jan./jun. 2014.

ALVES, M. T. G.; SOARES, J. F. O efeito das escolas no aprendizado dos alunos: um estudo com dados longitudinais no ensino fundamental. *Educação e Pesquisa*, São Paulo, v. 34, n. 3, p. 527-544, set./dez. 2008.

ALVES, M. T. G.; SOARES, J. F. Effects of schools and municipalities in the quality of basic education. *Cadernos de Pesquisa*, São Paulo, v. 43, n. 149, p. 492-517, maio/ago. 2013.

ANDRADE, R. J.; SOARES, J. F. O efeito da escola básica brasileira. *Estudos em Avaliação Educacional*, São Paulo, v. 19, n. 41, p. 379-406, set./dez. 2008.

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. *Teoria da resposta ao item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística, 2000.

AYALA, R. J. *The theory and practice of item response theory*. New York: The Guilford, 2009.

BATES, D.; MAECHLER, M.; BOLKER, B.; WALKER, S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, Innsbruck, Austria, v. 67, n. 1, p. 1-48, Oct. 2015.

BOTTAL, M.; ORSINI, N.; GERACI, M. A gradient search maximization algorithm for the asymmetric Laplace likelihood. *Journal of Statistical Computation and Simulation*, London, v. 85, n. 10, p. 1919-1925, 2015.

CHEN, Y. L.; TIAN, M. Z.; YU, K. M.; PAN, J. X. Composite hierarchical linear quantile regression. *Acta Mathematicae Applicatae Sinica*, London, v. 30, n. 1, p. 49-64, 2014.

DAVINO, C.; VISTOCCO, D. Quantile regression for the evaluation of student satisfaction. *Statistica Applicata*, London, v. 20, n. 3-4, p. 179-196, 2008.

GALARZA, C. E.; LACHOS, V. H.; BANDYOPADHYAY, D. Quantile regression in linear mixed models: a stochastic approximation EM approach. *Statistics and Its Interface*, London, v. 10, n. 3, p. 471-482, 2017.

GERACI, M. Linear quantile mixed models: the lqmm package for Laplace quantile regression. *Journal of Statistical Software*, Innsbruck, Austria, v. 57, n. 13, p. 1-29, Mayo 2014.

GERACI, M. *Package lqmm v. 1.5.3*. 2016. Disponível em: <<https://cran.r-project.org/web/packages/lqmm/lqmm.pdf>>. Acesso em: 5 set. 2017.

GERACI, M.; BOTTAL, M. Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics*, Oxford, UK, v. 8, n. 1, p. 140-154, 2007.

GERACI, M.; BOTTAL, M. Linear quantile mixed models. *Statistics and Computing*, London, v. 24, n. 3, p. 461-479, 2014.

GIMENES, N. Avaliação em larga escala no Brasil: tensões e desafios. *Estudos em Avaliação Educacional*, São Paulo, v. 26, n. 62, p. 254-261, maio/ago. 2015.

HASSAN, I. *Hierarchical quantile regression*. 2014. 34f. Thesis (Master of Science in Biostatistics) – University of Alberta, Canadá, 2014. Disponível em: <[https://era.library.ualberta.ca/files/m326m3245/Hassan\\_Imran\\_201408\\_MSc.pdf](https://era.library.ualberta.ca/files/m326m3245/Hassan_Imran_201408_MSc.pdf)>. Acesso em: 5 set. 2017.

KOENKER, R. *Quantile regression*. Cambridge, UK: Cambridge University, 2005.

KOENKER, R.; BASSETT, G. Regression quantiles. *Econometrica*, New York, v. 46, n. 1, p. 33-50, Jan. 1978.

LAROS, A. J.; MARCIANO, J. L.; ANDRADE, M. J. Fatores que afetam o desempenho na prova de Matemática do Saeb: um estudo multinível. *Avaliação Psicológica*, Porto Alegre, v. 9, n. 2, p. 173-186, ago. 2010.

MARTINS, P. S.; PEREIRA, P. T. Does education reduce wage inequality? Quantile regression evidence from 16 countries. *Labour Economics*, New York, v. 11, n. 3, p. 355-371, June 2004.

MORICONI, G. M.; NASCIMENTO, P. A. M. M. Fatores associados ao desempenho dos concluintes de engenharia no Enade 2011. *Estudos em Avaliação Educacional*, São Paulo, v. 25, n. 57, p. 248-278, jan./abr. 2014.

NASCIMENTO, P. A. M. M. Fatores associados ao desempenho escolar: resultado de um modelo multinível de valor adicionado. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 51, p. 16-39, jan./abr. 2012.

NETER, J.; KUTNER, M.; WASSERMAN, W.; NACHTSHEIM, C. *Applied linear statistical models*. Boston: McGraw-Hill/Irwin, 1996. 1415p.

PARÁ. Secretaria da Educação. *Estudos do SisPAE 2016 – Perfil dos participantes e fatores associados ao desempenho escolar*. Belém do Pará, fevereiro 2017. Disponível em: <<https://sispae.vunesp.com.br/Revistas2016.aspx>>. Acesso em: 7 mar. 2018.

PONTES; L. A. F.; SOARES, T. M. As metas escolares do Ideb: uma proposta alternativa de cálculo. *Estudos em Avaliação Educacional*, São Paulo, v. 27, n. 66, p. 690-715, set./dez. 2016.

RANGVID, B. S. Educational peer effects quantile regression evidence from Denmark with PISA 2000 data. *The economics of education and training*. Germany: Physica-Verlag, 2008.

R FOUNDATION FOR STATISTICAL COMPUTING. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2017. Disponível em: <[www.R-project.org](http://www.R-project.org)>. Acesso em: 5 set. 2017.

YU, K.; ZHANG, J. A three-parameter asymmetric Laplace distribution and its extension. *Communications in Statistics: Theory and Methods*, London, v. 34, n. 9/10, p. 1867-1879, Sept. 2006.

## APÊNDICE

As tabelas A1 a A5 apresentam os valores estimados dos coeficientes das regressões hierárquicas quantílicas, acompanhados dos respectivos erros-padrão e os valores-p do teste estatístico associado à hipótese nula de o coeficiente ser nulo, ou seja, não haver efeito significativo da referida variável ou do fator em análise.

Cabe observar que alguns coeficientes que não apresentaram significância estatística (valor-p maior que 0,05) não foram explicitados nas tabelas 3 e 4 do corpo do trabalho, porém nas tabelas A1 a A5 estão relacionados todos os resultados obtidos no ajuste dos modelos.

**TABELA A1 - Coeficientes, erros-padrão e valores-p para o quantil 10% - q(10%) do modelo de regressão quantílica hierárquica ajustado aos resultados do 8º ano do ensino fundamental do SisPAE 2016, por disciplina**

VARIÁVEIS/FATORES	LÍNGUA PORTUGUESA			MATEMÁTICA		
	ESTIMATIVAS	E.P.	VALOR-P	ESTIMATIVAS	E.P.	VALOR-P
Intercepto	145,8	1,4	<0,0001	168,9	1,0	<0,0001
<b>NÍVEL 1 - ESTUDANTE</b>						
Nível Socioeconômico da família do estudante (NSE)	2,5	0,3	<0,0001	1,9	0,2	<0,0001
Nível de Relacionamento Escolar estudante (NRE)	1,9	0,3	<0,0001	1,4	0,2	<0,0001
Nível do Clima Escolar estudante (NCE)	0,5	0,3	0,1022	0,1	0,2	0,8265
Sexo feminino	11,8	0,5	<0,0001	-0,9	0,3	0,0102
Estudante teve reprovação	-9,3	0,4	<0,0001	-6,2	0,4	<0,0001
Tempo gasto até a escola maior que 1/2 hora	-6,7	0,8	<0,0001	-3,7	0,5	<0,0001
Estudante aprende matéria com professor	8,2	0,8	<0,0001	4,2	0,6	<0,0001
Aprendizado depende mais do professor do que do estudante	-6,1	0,6	<0,0001	-3,5	0,5	<0,0001
Professor dá mais atenção aos estudantes bons	-5,2	0,5	<0,0001	-2,8	0,5	<0,0001
Professor se preocupa com o dever de casa	6,3	0,6	<0,0001	3,1	0,4	<0,0001
Professor costuma faltar	-5,3	0,5	<0,0001	-2,2	0,4	<0,0001
Estudante lê frequentemente	3,0	0,6	<0,0001	-	-	-
Pais incentivam a ler	3,3	0,7	<0,0001	-	-	-
<b>NÍVEL 2 - ESCOLA</b>						
Nível Socioeconômico do ano/escola	2,4	0,6	0,0001	1,9	0,4	<0,0001
Nível de Relacionamento Escolar da série/escola	1,1	0,6	0,1022	0,8	0,4	0,0489
Nível do Clima Escolar da série/escola	1,7	0,6	0,0078	1,6	0,4	<0,0001

Fonte: Elaboração dos autores, 2018.

**TABELA A2 - Coeficientes, erros-padrão e valores-p para o primeiro quartil - q(25%) do modelo de regressão quantílica hierárquica ajustado aos resultados do 8º ano do ensino fundamental do SisPAE 2016, por disciplina**

VARIÁVEIS/FATORES	LÍNGUA PORTUGUESA			MATEMÁTICA		
	ESTIMATIVAS	E.P.	VALOR-P	ESTIMATIVAS	E.P.	VALOR-P
Intercepto	168,2	2,8	<0,0001	187,9	1,0	<0,0001
<b>Disciplina: Língua Portuguesa</b>						
Nível Socioeconômico da família do estudante (NSE)	3,1	0,4	<0,0001	2,6	0,2	<0,0001
Nível de Relacionamento Escolar estudante (NRE)	2,0	0,3	<0,0001	1,4	0,2	<0,0001
Nível do Clima Escolar estudante (NCE)	0,4	0,4	0,2915	0,3	0,2	0,2206
Sexo feminino	13,0	0,7	<0,0001	-1,9	0,4	<0,0001
Estudante teve reprovação	-11,2	0,8	<0,0001	-8,0	0,4	<0,0001
Tempo gasto até a escola maior que 1/2 hora	-7,7	0,8	<0,0001	-4,5	0,5	<0,0001
Estudante aprende matéria com professor	7,1	0,9	<0,0001	5,3	0,5	<0,0001
Aprendizado depende mais do professor do que do estudante	-8,1	0,8	<0,0001	-4,9	0,4	<0,0001
Professor dá mais atenção aos estudantes bons	-6,1	0,9	<0,0001	-3,7	0,4	<0,0001
Professor se preocupa com o dever de casa	8,8	1,1	<0,0001	4,5	0,4	<0,0001
Professor costuma faltar	-5,5	1,0	<0,0001	-3,1	0,4	<0,0001
Estudante lê frequentemente	3,9	1,0	0,0002	-	-	-
Pais incentivam a ler	4,0	0,7	<0,0001	-	-	-
<b>Disciplina: Matemática</b>						
Nível Socioeconômico do ano/escola	3,2	1,2	0,0101	2,5	0,4	<0,0001
Nível de Relacionamento Escolar da série/escola	1,7	0,6	0,0049	1,3	0,5	0,0077
Nível do Clima Escolar da série/escola	2,5	0,8	0,0016	1,8	0,4	<0,0001

Fonte: Elaboração dos autores, 2018.

**TABELA A3 – Coeficientes, erros-padrão e valores-p para a mediana - q(50%) do modelo de regressão quantílica hierárquica ajustado aos resultados do 8º ano do ensino fundamental do SisPAE 2016, por disciplina**

VARIÁVEIS/FATORES	LÍNGUA PORTUGUESA			MATEMÁTICA		
	ESTIMATIVAS	E.P.	VALOR-P	ESTIMATIVAS	E.P.	VALOR-P
Intercepto	199,4	1,0	<0,0001	213,7	0,7	<0,0001
<b>NÍVEL 1 – ESTUDANTE</b>						
Nível Socioeconômico da família do estudante (NSE)	3,7	0,2	<0,0001	3,4	0,2	<0,0001
Nível de Relacionamento Escolar estudante (NRE)	2,0	0,2	<0,0001	1,5	0,2	<0,0001
Nível do Clima Escolar estudante (NCE)	0,6	0,2	0,0064	0,6	0,2	0,0016
Sexo feminino	11,7	0,4	<0,0001	-4,4	0,4	<0,0001
Estudante teve reprovação	-13,0	0,4	<0,0001	-10,3	0,4	<0,0001
Tempo gasto até a escola maior que 1/2 hora	-7,7	0,5	<0,0001	-4,9	0,5	<0,0001
Estudante aprende matéria com professor	5,4	0,6	<0,0001	6,0	0,5	<0,0001
Aprendizado depende mais do professor do que do estudante	-10,4	0,5	<0,0001	-6,9	0,4	<0,0001
Professor dá mais atenção aos estudantes bons	-6,4	0,5	<0,0001	-4,5	0,4	<0,0001
Professor se preocupa com o dever de casa	9,5	0,4	<0,0001	5,3	0,3	<0,0001
Professor costuma faltar	-5,5	0,5	<0,0001	-3,4	0,4	<0,0001
Estudante lê frequentemente	3,2	0,6	<0,0001	-	-	-
Pais incentivam a ler	5,3	0,5	<0,0001	-	-	-
<b>NÍVEL 2 – ESCOLA</b>						
Nível Socioeconômico do ano/escola	4,1	0,5	<0,0001	2,9	0,4	<0,0001
Nível de Relacionamento Escolar da série/escola	0,8	0,5	0,1129	1,2	0,5	0,0093
Nível do Clima Escolar da série/escola	2,3	0,3	<0,0001	2,2	0,4	<0,0001

Fonte: Elaboração dos autores, 2018.

**TABELA A4 - Coeficientes, erros-padrão e valores-p para o terceiro quartil - q(75%) do modelo de regressão quantílica hierárquica ajustado aos resultados do 8º ano do ensino fundamental do SisPAE 2016, por disciplina**

VARIÁVEIS/FATORES	LÍNGUA PORTUGUESA			MATEMÁTICA		
	ESTIMATIVAS	E.P.	VALOR-P	ESTIMATIVAS	E.P.	VALOR-P
Intercepto	227,3	1,6	<0,0001	240,7	1,2	<0,0001
<b>NÍVEL 1 - ESTUDANTE</b>						
Nível Socioeconômico da família do estudante (NSE)	5,1	0,2	<0,0001	4,3	0,2	<0,0001
Nível de Relacionamento Escolar estudante (NRE)	1,6	0,2	<0,0001	1,4	0,2	<0,0001
Nível do Clima Escolar estudante (NCE)	1,1	0,3	0,0001	1,0	0,2	<0,0001
Sexo feminino	10,4	0,4	<0,0001	-7,2	0,5	<0,0001
Estudante teve reprovação	-12,9	0,6	<0,0001	-12,2	0,4	<0,0001
Tempo gasto até a escola maior que 1/2 hora	-5,4	0,6	<0,0001	-5,0	0,6	<0,0001
Estudante aprende matéria com professor	2,9	0,7	0,0002	5,9	0,6	<0,0001
Aprendizado depende mais do professor do que do estudante	-11,7	0,5	<0,0001	-8,6	0,4	<0,0001
Professor dá mais atenção aos estudantes bons	-6,2	0,5	<0,0001	-5,0	0,5	<0,0001
Professor se preocupa com o dever de casa	11,1	0,6	<0,0001	5,5	0,5	<0,0001
Professor costuma faltar	-4,3	0,4	<0,0001	-4,1	0,5	<0,0001
Estudante lê frequentemente	4,9	0,5	<0,0001	-	-	-
Pais incentivam a ler	4,7	0,7	<0,0001	-	-	-
<b>NÍVEL 2 - ESCOLA</b>						
Nível Socioeconômico do ano/escola	4,7	0,3	<0,0001	2,7	0,4	<0,0001
Nível de Relacionamento Escolar da série/escola	1,6	0,4	0,0001	1,4	0,5	0,0039
Nível do Clima Escolar da série/escola	1,2	0,4	0,0033	2,7	0,4	<0,0001

Fonte: Elaboração dos autores, 2018.

**TABELA A5 – Coeficientes, erros-padrão e valores-p para o quantil 90% - q(90%) do modelo de regressão quantílica hierárquica ajustado aos resultados do 8º ano do ensino fundamental do SisPAE 2016, por disciplina**

VARIÁVEIS/FATORES	LÍNGUA PORTUGUESA			MATEMÁTICA		
	ESTIMATIVAS	E.P.	VALOR-P	ESTIMATIVAS	E.P.	VALOR-P
Intercepto	259,7	1,5	<0,0001	266,1	1,1	<0,0001
<b>NÍVEL 1 – ESTUDANTE</b>						
Nível Socioeconômico da família do estudante (NSE)	4,6	0,3	<0,0001	4,2	0,3	<0,0001
Nível de Relacionamento Escolar estudante (NRE)	1,3	0,3	0,0001	1,1	0,3	0,0001
Nível do Clima Escolar estudante (NCE)	0,6	0,4	0,0852	1,6	0,3	<0,0001
Sexo feminino	6,8	0,5	<0,0001	-9,6	0,5	<0,0001
Estudante teve reprovação	-15,1	0,7	<0,0001	-14,3	0,5	<0,0001
Tempo gasto até a escola maior que 1/2 hora	-5,4	0,8	<0,0001	-5,3	0,7	<0,0001
Estudante aprende matéria com professor	0,8	1,0	0,4084	5,9	0,7	<0,0001
Aprendizado depende mais do professor do que do estudante	-12,4	0,5	<0,0001	-9,5	0,5	<0,0001
Professor dá mais atenção aos estudantes bons	-7,4	0,6	<0,0001	-5,4	0,5	<0,0001
Professor se preocupa com o dever de casa	10,3	0,6	<0,0001	5,6	0,5	<0,0001
Professor costuma faltar	-4,9	0,6	<0,0001	-3,9	0,6	<0,0001
Estudante lê frequentemente	4,2	0,8	<0,0001	-	-	-
Pais incentivam a ler	3,8	0,7	<0,0001	-	-	-
<b>NÍVEL 2 – ESCOLA</b>						
Nível Socioeconômico do ano/escola	6,6	0,5	<0,0001	4,4	0,5	<0,0001
Nível de Relacionamento Escolar da série/escola	0,8	0,5	0,1680	0,3	0,6	0,6449
Nível do Clima Escolar da série/escola	1,8	0,7	0,0165	3,5	0,6	<0,0001

Fonte: Elaboração dos autores, 2018.

Recebido em: 20 SETEMBRO 2017

Aprovado para publicação em: 10 MAIO 2018



This content is licensed under a Creative Commons attribution-type BY-NC.