



UNIVERSIDADE DE BRASÍLIA – UnB
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, CONTABILIDADE E GESTÃO DE
POLÍTICAS PÚBLICAS – FACE
DEPARTAMENTO DE ECONOMIA – ECO
MESTRADO PROFISSIONAL EM ECONOMIA DO SETOR PÚBLICO – MESP

**OS EFEITOS DA FORMAÇÃO DOCENTE SOBRE O
DESEMPENHO ESCOLAR EM DIFERENTES PERCENTIS
DA DISTRIBUIÇÃO DE PROFICIÊNCIAS DO SAEB 2017**

HÉLIO JÚNIO ROCHA MORAIS

BRASÍLIA

2020

HÉLIO JÚNIO ROCHA MORAIS

**OS EFEITOS DA FORMAÇÃO DOCENTE SOBRE O DESEMPENHO
ESCOLAR EM DIFERENTES PERCENTIS DA DISTRIBUIÇÃO DE
PROFICIÊNCIAS DO SAEB 2017**

Dissertação apresentada à Faculdade de Economia, Administração e Contabilidade da Universidade de Brasília - UnB, como requisito parcial à obtenção do título de Mestre em Economia do Setor Público.

Orientadora: Profa. Dra. Maria Eduarda
Tannuri-Pianto

Brasília
2020

HÉLIO JÚNIO ROCHA MORAIS

**OS EFEITOS DA FORMAÇÃO DOCENTE SOBRE O DESEMPENHO ESCOLAR
EM DIFERENTES PERCENTIS DA DISTRIBUIÇÃO DE PROFICIÊNCIAS DO
SAEB 2017**

Dissertação apresentada à Faculdade de Economia, Administração e Contabilidade da Universidade de Brasília - UnB, como requisito parcial à obtenção do título de Mestre em Economia do Setor Público.

Aprovado em: ___/___/___

BANCA EXAMINADORA

Profa. Dra. Maria Eduarda Tannuri-Pianto (Orientadora)
Universidade de Brasília (Unb)

Profa. Dra. Ana Carolina Pereira Zoghbi
Universidade de Brasília (Unb)

Prof. Dr. Luís Gustavo de Amaral de Vinha
Universidade de Brasília (Unb)

À minha filha Mariana. Anseio que você tenha sucesso na sua trajetória escolar e profissional.

AGRADECIMENTOS

Agradeço a Deus pelo dom da vida. Nada seria possível sem a atuação dele na minha vida.

Agradeço ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) por disponibilizar os dados para realização da pesquisa, em especial ao Sr. Murillo Gameiro, Diretor de Gestão e Planejamento. Muitas pessoas contribuíram em diferentes etapas do projeto e registro meu agradecimento: Ao Professor Dr. Vander Lucas, Coordenador Acadêmico do Mestrado em Economia do Setor Público (MESP), que propiciou as condições necessárias para que eu tivesse a oportunidade de concluir o Mestrado. Às Professoras Dra. Maria Eduarda Tannuri-Pianto e Dra. Ana Carolina Pereira Zoghbi, que aceitaram conduzir a orientação dessa pesquisa. Muito obrigado pelos conselhos, apontamentos e o tempo dedicado neste período. Ao Pesquisador Geraldo Andrade, pelo auxílio e orientação no início da pesquisa. Ao Professor Chico Soares, que cedeu parte do seu tempo para trocar informações relevantes, as quais procurei incorporar ao texto. Ao Professor Dr. Wellington Tavares, que também compartilhou informações e um excelente material de pesquisa. Aos amigos do Inep, Denys Machado, Vanderlei Silva e Gizane Santos que, direta ou indiretamente, auxiliaram-me nesse percurso. E também à minha amiga Narah Rúbia, você também contribuiu com o resultado final.

Agradeço especialmente à minha esposa, Julliana, companheira sempre presente, pela compreensão e apoio incondicional durante essa longa caminhada.

RESUMO

Este estudo utiliza técnicas de regressão quantílica para identificar os efeitos da formação dos professores de matemática no desempenho dos estudantes da rede pública da educação básica brasileira. A abordagem quantílica permite analisar os efeitos das variáveis de interesse em diferentes pontos da distribuição de proficiência dos estudantes. Foram utilizados dados do Censo Escolar e do Sistema de Avaliação da Educação Básica (Saeb) para vincular os estudantes e os professores que lecionaram matemática durante a trajetória escolar dos anos finais do ensino fundamental. Os resultados foram apresentados em nível nacional e também por região e redes de ensino. Em diferentes recortes foi possível identificar que as características da formação dos professores afetam o desempenho dos estudantes.

Palavras-chave: Formação do professor, Regressão quantílica, Desempenho de estudantes.

ABSTRACT

This study uses quantile regression techniques to identify the effects of mathematics teachers training on the students performance in the public system of Brazilian basic education. The quantile approach allows to analyze the effects of variables of interest at different points of the students proficiency distribution. School Census and Basic Education Assessment System (Saeb) data have been used to link students and teachers who taught mathematics during the school trajectory of the final years of elementary school. The results have been presented at the national level and also by region and education networks. In different groups, it was possible to identify which teachers training characteristics have an effect on the students' performance.

Keywords: Teacher training, Quantile regression, Students' performance.

LISTA DE QUADROS E TABELAS

Quadro 1 – Descrição das variáveis utilizadas no modelo	35
Quadro 2 – Questões 39 e 40 do Questionário do Diretor – Saeb 2017	43
Tabela 1 - Número de estudantes previstos no Saeb 2017	25
Tabela 2 – Número de estudantes que participaram do Saeb 2017 por etapa de ensino	26
Tabela 3 – Estudantes com notas válidas em 2017 e 2013.....	26
Tabela 4 – Estudantes, turmas, escolas e professores identificados no período de 2017-2013	27
Tabela 5 – Estudantes quem compõem a base de dados final	27
Tabela 6 – Professores com formação adequada em matemática	29
Tabela 7 – Características dos estudantes, professores e escolas	30
Tabela 8 – Modelo apenas com o intercepto (modelo nulo)	32
Tabela 9 – Modelo com o intercepto e a adequação da formação do professor.....	33
Tabela 10 – Modelo com o intercepto, proficiência anterior e adequação da formação do professor	34
Tabela 11 – Modelo Brasil	36
Tabela 12 – Proficiência dos estudantes em matemática do 9º ano do Ensino Fundamental no Saeb (2017).....	38
Tabela 13 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, no Brasil.	39
Tabela 14 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, no Brasil.	44

SUMÁRIO

1. INTRODUÇÃO	9
2. REVISÃO DA LITERATURA	11
2.1 Sistema de avaliação educacional no Brasil	15
3. METODOLOGIA DE ANÁLISE	18
3.1 O modelo de regressão quantílica	18
3.2 Estrutura hierárquica	22
3.3 O modelo de regressão quantílico hierárquico	23
4. BASE DE DADOS	25
5. RESULTADOS E DISCUSSÃO	31
6. CONSIDERAÇÕES FINAIS.....	45
7. BIBLIOGRAFIA	47
8. ANEXO.....	52

1. INTRODUÇÃO

A formulação e o desenvolvimento de políticas públicas educacionais recebem especial atenção na agenda governamental. A importância dessas políticas pode estar associada a vários fatores, dentre os quais destacam-se os benefícios e os custos de investimentos que resultam em melhorias na qualidade de escolas que, em sua grande maioria, são responsáveis pelo desenvolvimento de competências relevantes para o mercado de trabalho.

No Brasil, a prestação de serviços educacionais pelo poder público compreende a oferta de educação básica, que é uma obrigação do Estado. Esses serviços são fundamentados em exigências constitucionais e visam oferecer um atendimento equitativo e qualitativo.

As políticas públicas de educação nos países em desenvolvimento, como o Brasil, concentram-se, em grande medida, no aumento de matrículas. O intuito é garantir a universalização do acesso à educação básica. Como resultado das políticas desenvolvidas pelo Ministério da Educação nos últimos 20 anos, os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2018 confirmam aumento substancial na taxa de matrícula e apontam que 99,3% das crianças de 6 a 14 anos estão matriculadas no ensino fundamental.

No entanto, com o aumento da taxa de cobertura, surge um desafio ainda maior: melhorar os níveis de aprendizagem nas escolas. Os relatórios que apresentam os resultados das escolas brasileiras nas avaliações nacionais de larga escala demonstram baixo desempenho dos estudantes, especialmente nas áreas de língua portuguesa e matemática. O mesmo ocorre quando os resultados das escolas brasileiras são comparados aos de países desenvolvidos nas avaliações realizadas pelo Program for International Student Assessment (PISA).

Na tentativa de melhorar a qualidade do ensino, convencionalmente, políticas educacionais são direcionadas no sentido de aumentar os recursos destinados à educação, fornecendo livros didáticos, melhorando a infraestrutura escolar, capacitando professores e adequando o número de alunos por sala. Essas são algumas das ações que podem influenciar positivamente a aprendizagem dos alunos.

Em outras palavras, identificar essas variáveis pode ajudar a gerar resultados positivos no desempenho dos alunos. Nesse contexto, os professores exercem papel de destaque. A eficácia do professor pode ser medida a partir do desempenho dos alunos, avaliado por meio de testes cognitivos. Essa abordagem permite mensurar e analisar as características dos professores que mais contribuem com o sucesso escolar. Nesse caso, espera-se que professores mais

experientes, com maior titulação e com formação adequada na área que leciona sejam mais efetivos, impactando positivamente o desempenho escolar.

Este estudo analisa especificamente os efeitos da formação docente sobre o desempenho escolar em matemática nos anos finais do ensino fundamental de escolas públicas brasileiras nos exames de avaliação nacional aplicados em 2017 e buscar responder ao seguinte problema de pesquisa: a formação adequada dos professores afeta o desempenho dos estudantes brasileiros nas avaliações externas de larga escala?

Nesse estudo, a formação do professor será adequada quando o docente que leciona matemática no ensino fundamental tiver formação superior na respectiva disciplina, seja licenciatura, ou bacharelado com curso de complementação pedagógica concluído.

Diante do exposto, este estudo tem como objetivo geral, mensurar qual é o impacto da formação dos professores que lecionam matemática no desempenho de estudantes de escolas públicas ao longo do ensino fundamental; e ainda, como objetivos específicos: identificar se o efeito da formação do professor é maior em estudantes de alto desempenho ou de baixo desempenho; e comparar o impacto dessas variáveis nas diferentes redes de ensino (estadual e municipal) e nas cinco regiões do país.

2. REVISÃO DA LITERATURA

Desde os anos 60, com os primeiros estudos sobre os determinantes da função de produção de educação, as conclusões sobre a importância dos fatores educacionais nos resultados de ensino são controversas (Coleman, 1966). Especialmente no caso dos países em desenvolvimento, esses primeiros estudos não mostram uma relação positiva entre os resultados empíricos e a teoria convencional dos determinantes que afetam o desempenho escolar dos alunos como, por exemplo, as condições de infraestrutura escolar, a formação de professores, o tamanho das turmas, e o nível de despesas por aluno (Heyneman, 1986; Hanushek, 1992). Isso levou Hanushek a afirmar, em vários artigos, que não há uma relação direta entre o desempenho dos alunos e a quantidade de recursos alocados por escola. Após três décadas de estudos, Hanushek (2006) chegou à conclusão que variações de recursos escolares não estão relacionadas à variações nos desempenhos dos alunos.

De acordo com Coleman (1966), a escola teria relativamente pouca importância na determinação do desempenho do aluno. Na verdade, segundo o autor, os fatores de sucesso do desempenho escolar estavam mais associados aos próprios alunos, aos pais e a outros fatores combinados.

Por outro lado, alguns autores criticam esse pensamento, pois identificam problemas metodológicos inerentes aos estudos em questão (Krueger, 1999). Segundo especialistas, esses problemas metodológicos podem ser equacionados utilizando experimentos aleatórios (Kremer, 2003; Glewwe e Kremer, 2006). Utilizando essa metodologia, estudos realizados nos Estados Unidos e em países em desenvolvimento mostraram que fatores como o tamanho das turmas estão, sim, associados ao desempenho dos alunos (Krueger, 1999).

Alguns estudos destacam o papel de incentivos na educação (Kremer, 2003; Hanushek, 1955). Uma série de pesquisas investigaram a efetividade de várias reformas educacionais, incluindo a remuneração dos professores por desempenho (Glewwe, 2003; Rivkin et al (2005). Esses estudos sugerem que a política de incentivos pode render ganhos substanciais nos resultados acadêmicos. Isso porque um importante determinante de rendimento escolar dos estudantes é o esforço do professor em proporcionar as condições necessárias para o melhor aproveitamento das aulas.

Em relação às características dos professores, muitos pesquisadores (HANUSHEK, 1986; FIGLIO, 1997; ARMOR et al., 1975; RIVKIN et al., 2003; AARONSON et al., 2007; ROCKOFF, 2005; RIVKIN et al., 2001; HANUSHEK et al., 2004; SUMMERS e WOLFE,

1977; HANUSHEK, 1971; RIVKIN et al., 2005; COOPER e COHN, 1997; EHRENBERG e BREWER, 1994; e JACOB e LEFGREN, 2002; WOESSMANN, 2003) concluíram que, embora qualificação represente um componente importante para o processo de aprendizagem, outras variáveis afetam mais o desempenho dos alunos. Além disso, a qualificação do professor é, em geral, difícil de ser mensurada diretamente.

Hanushek (1986) estimou os efeitos das características dos professores sobre o desempenho dos estudantes e não identificou evidências consistentes de que professores mais qualificados tiveram um impacto mais positivo nas notas dos estudantes.

Para Figlio (1997), estudos que buscam medir a relação entre as características dos professores e o desempenho dos alunos podem gerar resultados enviesados, pois professores mais qualificados podem atuar em escolas e municípios com alunos de melhor desempenho. Nessa linha, Kane e Staiger (2008) e Kane et al.(2013) também reportaram resultados enviesados, mesmo nos estudos em que os alunos foram atribuídos aos professores de forma aleatória.

Por outro lado, os estudos apresentados por Armor et al. (1975), com dados coletados diretamente nas unidades escolares, e os estudos realizados por Rivkin et al. (2003) e Aaronson et al. (2007), utilizando bases de dados proveniente de instituições de pesquisa, constataram que o desempenho do aluno é afetado pela qualidade de seus professores. Esses pesquisadores reportaram que as dificuldades empíricas nesse tipo de estudo estão relacionadas à qualidade dos dados.

Rockoff (2005) mostra que a qualidade dos professores aumenta os resultados dos testes em aproximadamente 0,1 desvios-padrão em leitura e em matemática nos testes aplicados nacionalmente. Os anos de experiência do professor, por exemplo, representam um aumento de 0,17 desvios-padrão na nota dos alunos, quando os professores de início de carreira são comparados com aqueles com mais dez anos de experiência.

Os resultados obtidos por Rockoff (2005) reforçam a tese de que investir na qualidade dos professores é um caminho importante para melhorar o desempenho educacional dos estudantes. O autor adverte, por outro lado, que políticas voltadas à qualificação dos professores devem objetivar sua formação e não o desempenho dos estudantes, sob o risco de se criar uma anomalia nas unidades escolares, porque a distribuição dos professores, em geral, pode levar em consideração o desempenho prévio dos estudantes. Outra ressalva feita pelo autor é a de que, é possível, a partir de um determinado período, que a experiência adicional dos professores não afete o desempenho dos estudantes.

Nessa mesma linha, Rivkin et al. (2001) afirmam que o efeito marginal de anos de experiência do professor diminui rapidamente e que os efeitos positivos da experiência ocorrem nos primeiros anos. Além disso, por vezes, aos professores com mais tempo de atuação na escola, pode ser facultada a escolha da turma que irá lecionar, e eles podem optar pelas turmas cujos alunos tenham melhor desempenho, afetando os resultados alcançados por Hanushek et al. (2004).

Desse modo, é possível compreender que as características dos professores, incluindo anos de experiência e formação, têm pouco impacto sobre a variação de desempenho dos estudantes.

Summers e Wolfe (1977) destacam que a utilização de dados desagregados, isto é, dados individuais, implica resultados mais precisos, sendo a forma mais adequada para medir o efeito no desempenho educacional a partir de variáveis atribuídas diretamente ao poder público (professor, infraestrutura escolar, currículo).

Summers e Wolfe (1977) ainda dedicam especial atenção aos professores e buscam identificar as características associadas a eles que mais afetam o desempenho dos estudantes. Os autores identificam evidências de que o desempenho dos melhores alunos está positivamente relacionado ao tipo de graduação do professor, com maior feito sobre os estudantes de menor poder aquisitivo. Nesse sentido, espera-se que professores mais qualificados e com maior grau de instrução afetem o desempenho escolar.

Contudo, os resultados encontrados em relação à experiência dos professores mostraram-se irrelevantes, corroborando a pesquisa de Hanushek (1971). Em relação à experiência dos professores, estudos mais recentes mostram que há uma relação positiva entre desempenho do aluno e essa variável. Exemplos são os estudos de Goldhaber e Anthony (2007), Jepsen (2005), Krueger, (1999), Rivkin et al. (2005), Rochoff (2004) e Kukla-Acevedo (2009). Poucos estudos reportaram estimativas estatisticamente significantes, como é o caso da pesquisa conduzida por Hanushek e Rivkin (2006).

Rivkin et al. (2005) apontam que a qualidade dos professores é uma variável importante e que influencia o desempenho dos estudantes. Esse impacto, entretanto, afeta estudantes de maneira diferente, isto é, a eficácia dos professores não gera o mesmo efeito em todos os alunos. Para estes autores, professores bem qualificados afetam positivamente o desempenho especificamente dos estudantes de matemática, superando inclusive a desvantagem dos estudantes com baixo poder aquisitivo. Em relação aos títulos obtidos, os

autores não encontram evidências de que professores com cursos de pós-graduação *stricto sensu* afetem positivamente o desempenho dos estudantes.

Cooper e Cohn (1997), Ehrenberg e Brewer (1994), e Jacob e Lefgren (2002) também não identificaram efeitos positivos no desempenho dos estudantes em função da formação dos professores.

Entretanto, utilizando dados do Trends in Mathematics and Science Study (TIMSS)¹ de 1995, Woessmann (2003) constatou que a formação dos professores afeta positivamente o desempenho dos alunos. Dos nove países europeus pesquisados, os estudantes noruegueses tiveram o maior efeito positivo e os portugueses registram menor relação entre as variáveis de formação dos professores e desempenho dos estudantes.

Os resultados do estudo mostram também que a experiência de um professor iniciante de matemática, ou seja, com um ano de atuação, afeta negativamente o desempenho do estudante em 2,1 pontos quando comparado ao professor com dois anos de experiência. A diferença aumenta para 6,9 pontos quando a comparação é feita com um professor experiente, isto é, com 10 anos de atuação.

Campos (2014) ensina que o desenvolvimento cognitivo é um processo cumulativo, e por isso a análise de fatores que buscam explicar as variáveis que afetam o desempenho dos estudantes na sua trajetória escolar, deve considerar os diferentes insumos que possam exercer influência na função de produção da educação, como por exemplo, as variáveis escolares, as características individuais dos estudantes e familiares.

Na prática, dada a complexidade e impossibilidade de se incorporar ao modelo econométrico todas as variáveis que afetam o desempenho dos estudantes, pode ocorrer viés por omissão de variáveis relevantes no modelo.

Chetty, Friedman e Rockof (2014) afirmam que grande parte das variáveis omitidas que afetam o desempenho dos estudantes pode ser capturada pela nota de testes anteriores. Em sintonia com Heckman, Ichimura e Todd (1997) e Heckman et al. (1998), os pesquisadores apontam que é possível contornar os problemas decorrentes de variáveis omitidas que geram estimativas enviesadas quando as notas com o desempenho anterior dos estudantes são incluídas nos modelos econométricos. Nesse sentido, afirmam os autores que os modelos que não adotam esse controle produzem vieses superiores a 40%.

¹ O Trends in Mathematics and Science Study (TIMSS) é uma avaliação internacional de desempenho dos estudantes em matemática e em ciências, realizado pela International Association for the Evaluation of Student Achievement (IEA).

A escolha deste estudo por estimar os resultados utilizando apenas os resultados dos testes em matemática, mesmo tendo à disposição os resultados dos testes de leitura, não é aleatória. Segundo Daniel Aaronson (2007), a ligação entre os professores de matemática e os resultados dos testes de matemática são mais precisos do que para qualquer outra disciplina. Além disso, segundo o pesquisador, os resultados dos testes de matemática parecem ter mais poder preditivo do que os resultados dos testes de leitura. Isso é o que outros estudos, como, por exemplo, Grogger e Eide (1995); e Hanushek e Kimko (2000), frequentemente afirmam.

2.1 Sistema de avaliação educacional no Brasil

A avaliação do sistema educacional brasileiro é de responsabilidade do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), autarquia federal vinculada ao Ministério da Educação (MEC). A metodologia de avaliação está estruturada no Sistema de Avaliação da Educação Básica (Saeb), que compreende a avaliação da educação básica separada pelas etapas de ensino. Atualmente, a avaliação ocorre a cada dois anos, sempre no final de cada etapa do ciclo de ensino: 5º ano (4ª série) e 9º ano (8ª série) do Ensino Fundamental e 3ª série do Ensino Médio. O Saeb teve início em 1990, quando a avaliação ainda não era realizada no final de cada ciclo. Participaram das primeiras edições alunos da 1ª e 3ª série, referente aos anos iniciais, e da 5ª e 7ª séries dos anos finais do Ensino Fundamental.

Trata-se, portanto, de uma avaliação em larga escala realizada em âmbito nacional, com vistas a mensurar o desempenho dos alunos nos testes padronizados. Para tanto, são definidas amostras consistentes e representativas de vários estratos do universo (todos os alunos da rede de ensino), contemplando todos os Estados e o Distrito Federal. Além dos testes propriamente ditos, são coletadas informações intraclasse, extraclasse e socioeconômicas dos alunos, viabilizando a observação e associação entre fatores determinantes no desempenho escolar. Informações sobre a infraestrutura das unidades de ensino, sobre o perfil dos gestores educacionais e sobre as práticas pedagógicas dos professores também fazem parte dos questionários do Saeb.

Dessa forma, o conjunto de informações do Saeb ao longo do tempo permite um acompanhamento sofisticado da evolução do processo de aprendizagem e dos fatores associados à qualidade da educação. Esses resultados são utilizados na elaboração e na definição das políticas públicas educacionais do Ministério da Educação e, ao mesmo tempo, demonstram aos gestores educacionais, professores, pais e alunos a situação da educação

brasileira. Os resultados sinalizam também, especialmente à comunidade escolar, as distorções e a necessidade de intervenções para o aperfeiçoamento do processo de ensino e de aprendizagem dos alunos.

Desde a sua criação em 1990, o Saeb foi reformulado continuamente, ora para atender demandas da sociedade acadêmica, especialistas e pesquisadores, ora para atender sua própria característica de retratar com fidedignidade os diversos perfis da educação brasileira. Foram realizadas mudanças metodológicas, operacionais e de abrangência. Na primeira edição, foram avaliadas escolas do Ensino Fundamental da rede pública urbana, nas disciplinas de língua portuguesa, ciências e matemática.

Mudanças nos procedimentos foram implementadas a partir da edição de 1995, quando foram avaliadas as turmas do final de cada ciclo. Nessa mesma edição, os testes foram elaborados com base na moderna Teoria de Resposta ao Item (TRI), em substituição à teoria clássica dos testes. A TRI é um paradigma para desenho, análise e pontuação de testes no qual a resposta a um item é modelada de acordo com a função de proficiência do aluno e com os parâmetros que expressam certas propriedades do item. Quanto maior a proficiência, maior a probabilidade de o aluno acertar o item (Klein, 1995).

Outra alteração importante no Saeb a partir de 2005 foi a criação da Prova Brasil. A nova avaliação seguiu o mesmo formato da anterior, utilizando a TRI e as matrizes de referências do Saeb. A diferença crucial refere-se à abrangência, pois todas as escolas da rede pública da 4ª e 8ª série do Ensino Fundamental com pelo menos 30 alunos localizadas na zona urbana participaram da avaliação. Com isto, a Prova Brasil permitiu mensurar a proficiência média de cada estabelecimento de ensino, município, unidade da federação e Brasil. Por sua vez, o Saeb, por se tratar de uma amostra, gerava resultados apenas para os estratos de interesse.

Outra característica da Prova Brasil é que todos os alunos são submetidos aos testes de língua portuguesa e matemática. No Saeb, havia sorteio de escolas e depois de turmas, sendo que apenas metade da turma respondia questões de matemática e outra metade de língua portuguesa. Como a Prova Brasil permite a análise de desempenho por município e por escolas, soluciona o problema apontado por Cotta (2001): o de que os dados do Saeb eram limitados e não captavam as influências locais existentes no âmbito de cada unidade escolar.

Participaram da primeira edição da Prova Brasil 3.392.880 da 4ª e da 8ª séries do Ensino Fundamental em 5.387 municípios, num total de 40.962 e 125.852 turmas (Inep, 2007). Em 2007, ocorreu uma nova ampliação no público da Prova Brasil e do Saeb. O critério de

seleção das escolas, que até 2005 era de 25 alunos por escola, foi alterado para escolas com pelo menos 20 alunos matriculados.

Embora a Prova Brasil tenha caráter censitário, há restrições quanto às outras etapas de ensino não contempladas no universo da avaliação. Por isso, o Saeb continua sendo realizado de forma complementar à Prova Brasil. Como a metodologia utilizada nas duas avaliações é a mesma, a aplicação ocorre no mesmo período. O conjunto de informações obtidas pelo Saeb e pela Prova Brasil, associados aos dados do Censo Escolar, constitui para a sociedade em geral, a comunidade escolar, o MEC e para as secretarias estaduais e municipais de educação uma visão mais clara e precisa das ações que precisam ser adotadas para melhorar o nível da qualidade da educação brasileira, em especial, da educação pública, além de apontar áreas prioritárias para diminuição das desigualdades existentes entre as redes de ensino.

Uma crítica ao novo modelo de avaliação concentra-se na complexidade com a qual alguns dados são apresentados. Os resultados do Saeb, por exemplo, devem ser interpretados a partir da escala de proficiência que está dividida em vários níveis e apontam os conhecimentos e habilidades que os alunos demonstram. Para Franco (2001) é importante desenvolver mecanismos que facilitem a interpretação desses resultados por meio de gráficos simples, capazes de traduzir ao público em geral a modelagem complexa que existe por trás dos modelos estatísticos utilizados na elaboração do Saeb.

Neste mesmo pensamento, Cotta (2001) defende que usualmente os resultados das avaliações não correspondem às expectativas da sociedade simplesmente pela presunção de que os conhecimentos necessários para interpretação e utilização dos resultados já são dados. Para a autora, é necessário criar condições específicas que facilitem a compreensão dos resultados sob o risco de comprometer o esforço empenhado na realização da avaliação.

Outro estudo realizado sobre a utilização dos dados do Saeb feito por Horta Neto (2007) aponta as resistências existentes nas escolas em discutir o desempenho dos alunos a partir das informações coletadas por meio do sistema de avaliação. Entre as dificuldades estão o caráter político, já que algumas escolas entendem o sistema como um instrumento imposto pelo Governo para controle, e não uma ferramenta de diagnóstico que permite identificar as habilidades alcançadas e aquelas que ainda precisam ser desenvolvidas.

Torna-se clara a necessidade de construir um elo entre as avaliações de larga escala e a realidade vivida por professores e alunos nas milhares de escolas espalhadas por todo o Brasil, a fim de que os exames possam constantemente aperfeiçoar os mecanismos utilizados para

retratar com mais exatidão os aspectos que interferem diretamente no sistema educacional brasileiro, bem como o nível de desempenho dos alunos que fazem parte desse sistema.

3. METODOLOGIA DE ANÁLISE

3.1 O modelo de regressão quantílica

Os estudos que analisam os fatores associados ao desempenho dos estudantes, usualmente, recorrem a modelos de regressão por Mínimos Quadrados Ordinários (MQO), que estimam o efeito das variáveis intraescolares e extraescolares no desempenho dos estudantes médios (Neter et al., 1996).

O modelo de regressão clássico apresenta resultados para média condicional da variável de interesse associada a uma alteração nas covariáveis. No entanto, essa metodologia padrão pode não ser a mais adequada para avaliação de diversas políticas educacionais, uma vez que não demonstra como as características escolares afetam o aprendizado dos estudantes nos diferentes pontos da distribuição da escala dos testes.

Por exemplo, um determinado estudo sobre o impacto que aumento de investimento público por estudante pode exercer na qualidade da educação pode identificar os efeitos positivos dessa política para as notas médias do teste. Entretanto, seria útil diferenciar esses efeitos para os estudantes que estão na parte inferior da distribuição condicional. A abordagem mais completa, além de identificar se a política educacional avaliada é relevante, indicaria o grupo que é mais afetado pela política. Isso é muito importante para formulação de políticas mais focadas.

Enquanto o modelo clássico de regressão linear está baseado na mudança da média condicional da variável dependente, o modelo de regressão quantílica especifica mudanças nos percentis condicionais. Dessa forma, como diferentes percentis podem ser analisados, é possível obter resultados de como as variáveis explicativas afetam a variável dependente nos diferentes pontos da distribuição condicional obtendo informações sobre mudanças na localização, disseminação e forma, como mostram Koenker (2005) e Davino et al. (2013).

Para analisar como uma determinada variável afeta mais fortemente os estudantes da parte inferior ou superior na escala de proficiência, pode-se adotar o modelo de Regressão Quantílica proposto por Koenker e Basset (1978).

As regressões quantílicas propostas por Koenker e Basset (1978) são uma metodologia de regressão “robusta”, nos casos de não normalidade dos resíduos, isto é, quando os erros não

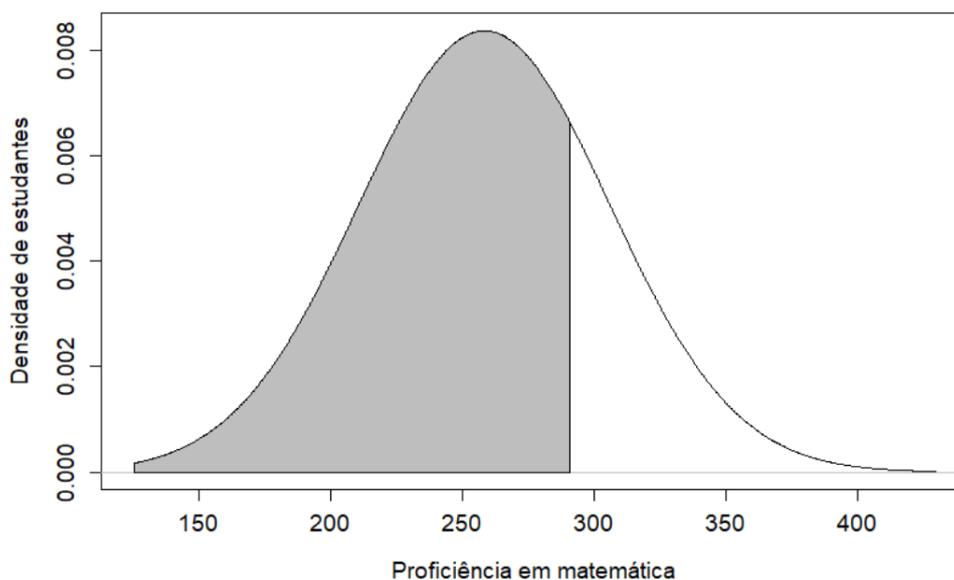
são normalmente distribuídos, quando os erros são heterocedásticos ou quando existe forte assimetria no comportamento da variável resposta Chen et al. (2014)

Essa técnica de regressão quantílica tem sido utilizada não apenas por essa propriedade robusta, mas também por permitir ao analista estimar efeitos em diferentes pontos da distribuição condicional dos resultados. O segundo quartil (Q_2) separa os 50% menores valores dos 50% maiores valores, e coincide com a mediana.

Neste estudo, será denominado $q(k)$ o quantil dos estudantes avaliados nos diferentes pontos da escala de proficiência do Saeb, isto é, o valor que separa os $k\%$ menores valores dos maiores valores.

A Figura 1 ilustra a posição do terceiro quartil incondicional (Q_3) denotado por $q(0,75)$, que deixa os 75% dos estudantes abaixo dele.

Figura 1 – Ilustração do terceiro quartil incondicional: ponto que separa os 75% de estudantes com menor proficiência dos 25% com maior proficiência.



Fonte: Elaboração própria.

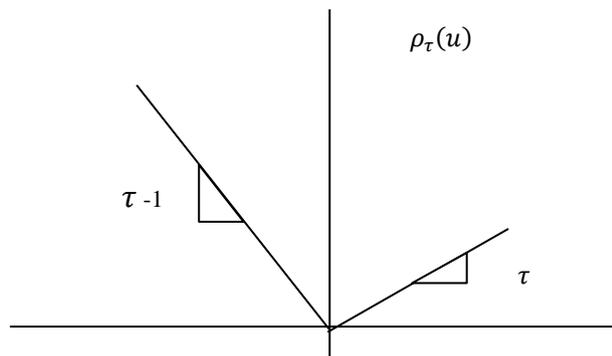
A regressão quantílica tem algumas vantagens em relação ao modelo de regressão linear clássico. Além de permitir a análise nos efeitos no quantil e não apenas na média, no caso da mediana, o estimador da regressão quantílica é menos sensível às observações extremas (*outliers*) presentes na variável dependente, do que no modelo de Mínimos Quadrados Ordinários, diminuindo o impacto dessas observações sobre os resultados.

O modelo utilizado neste trabalho baseia-se na regressão quantílica proposta por Koenker e Bassett (1978). Esse modelo possibilita medir o efeito da variável resposta de forma mais completa em diferentes quantis da distribuição condicional.

Formalmente, uma variável aleatória Y pode ser caracterizada por sua função de distribuição, $F(Y) = Prob(Y \leq y)$ para quaisquer $0 < \tau < 1$; em que, $Q(\tau) = \inf \{y: F(Y) \geq \tau\}$ denomina o τ -ésimo percentil de Y . A mediana $\tau = 0,5$, representa a posição central da distribuição. Assim como uma função de probabilidades, a função quantílica fornece uma caracterização completa da variável aleatória Y .

Os percentis podem ser formulados a partir da resolução de um simples problema de otimização. Para quaisquer $0 < \tau < 1$, define-se a função perda "check function", $\rho_\tau(u) = u(\tau - I(u > 0))$, representada na figura 2.

Figura 2. Função perda da regressão quantílica



Fonte: Koenker (2005).

De outro modo, dada uma amostra de n observações da variável Y , o quantil amostral de ordem τ é a solução para a função objetivo que deve ser minimizada:

$$\min_{q \in \mathbb{R}} \sum_{i=1}^n \rho_\tau(y_i - q) \tag{1}$$

Koenker (2005) mostra como calcular o quantil condicional de Y , dado x como uma função linear nos parâmetros na forma $Q_\tau(Y|x) = x'\beta(\tau)$, onde $\beta(\tau)$ é um vetor de parâmetros para o quantil condicional τ .

Para estimar $\beta(\tau)$ deve-se encontrar $\hat{\beta}(\tau)$ que seja a solução do problema de otimização:

$$\min_{\beta \in \mathbb{R}} \sum_{i=1}^n \rho_\tau(y_i - x_i'\beta) \quad (2)$$

A função ρ_τ multiplica os resíduos por τ se eles forem não-negativos e por $(\tau - 1)$ caso contrário, conforme apresentado na figura 2.

Assim, com base em um conjunto de valores (y_1, y_2, \dots, y_n) , pelo método dos mínimos quadrados ordinários, se:

$$y_i = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_{ip} + \varepsilon_1,$$

Onde ε_1 tem média igual a zero e a média condicional de $Y|X$ pode ser escrita como:

$$E(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_{ip} \quad E(y|x)$$

Então, a distribuição condicional da variável Y , supondo as relações lineares:

$$y_i = \beta_0(\tau) + \beta_1(\tau)x_1 + \dots + \beta_p(\tau)x_{ip} + u_i,$$

Onde u_1 são variáveis aleatórias independentes. O quantil da ordem τ de $Y|X$ pode ser escrito como:

$$Q_\tau(x) = \beta_0(\tau) + \beta_1(\tau)x_1 + \dots + \beta_p(\tau)x_p$$

3.2 Estrutura hierárquica

Segundo Hanushek (1979), os desempenhos dos estudantes são afetados por um conjunto de fatores educacionais: família, características inatas do indivíduo, variáveis da escola e da comunidade local e, ainda, a influência dos pares.

Os efeitos no desempenho dos estudantes provocados por cada grupo de variáveis não são uniformes. Um determinado grupo de variáveis afeta o lado da demanda por educação, e outro afeta a oferta educacional.

No Brasil, é comum que estudos que exploram os efeitos no desempenho dos estudantes incluam no mesmo nível variáveis escolares e da comunidade local, e de características individuais e da família por meio de regressões pelo método dos Mínimos Quadrados Ordinários (MQO). Sobre essa abordagem, tem-se que, quando todas as informações são analisadas apenas no nível agregado, não é possível identificar as características individuais que também afetam a variável de interesse, ocasionando resultados imprecisos.

De maneira similar, os modelos construídos apenas com variáveis individuais geram resultados que, na verdade, carregam informações do nível agregado, isto é, do grupo ao qual o indivíduo faz parte.

A modelagem multinível ou hierárquica possibilita ao pesquisador controlar as variáveis de acordo com suas características de agrupamento, isto é, leva em consideração a estrutura aninhada dos registros observados. A regressão hierárquica é uma extensão do modelo de regressão linear clássico, podendo ser aplicada em uma ampla variedade de estudos.

As primeiras pesquisas com aplicação de modelos hierárquicos foram desenvolvidas em meados da década de 1980, e apresentadas em artigos científicos por diferentes autores, Mason et al., 1983; Aitkin e Longford, 1986; Leeuw e Kreft, 1986; Goldstein, 1987, 1995; e Bryk e Raudenbush, 1992. Os dois principais grupos de pesquisa em estatísticas educacionais conduzidos, respectivamente, por Goldstein e Raudenbush desenvolveram os principais softwares utilizados na modelagem multinível. O programa HLM por (Raudenbush & Bryk, 2002) e o MLwin por (Goldstein, 1995). Esse referencial teórico, os softwares desenvolvidos e os estudos subsequentes, como Snijders e Bosker, 1999; e Hox, 2002, foram fundamentais para a evolução e a consolidação desse modelo de análise.

Para Soares (2007) a análise multinível tornou-se a referência para qualquer pesquisa que utilize dados em larga escala com diferentes níveis de agregação. Esse método tem sido

aplicado nos estudos educacionais, porque possibilita que se estabeleça essa hierarquia, ou níveis na estrutura dos dados educacionais. Em um nível estão as informações de contexto escolar e, em outro, as informações individuais dos estudantes.

As vantagens da regressão multinível, quando comparadas ao modelo de regressão linear clássico, são apresentadas por Ferrão, Beltrão e Santos (2002), ao destacarem que a modelagem de dados educacionais pela modelo linear clássico resulta em estimativas do erro padrão muito pequenas, o que ocorre quando existe forte correlação intraclasse. Por outro lado, as estimativas obtidas no modelo multinível geram resultados mais conservadores. Além disso, no modelo de regressão multinível o pesquisador analisa a variância do erro em cada nível hierárquico, permitindo assim, melhor interpretação dos resultados, distinguindo, por exemplo, os efeitos das variáveis intraescolares das variáveis extraescolares.

3.3 O modelo de regressão quantílico hierárquico

O processo para estimar os coeficientes no modelo quantílico hierárquico (ou multinível) é semelhante ao modelo hierárquico linear (HLM). A principal diferença entre as duas abordagens é que no modelo tradicional as regressões estão em torno da média, enquanto as condições de regressão no modelo quantílico ocorrem em diferentes pontos de interesse do pesquisador.

As estimações do modelo de regressão quantílica hierárquica podem ser realizadas utilizando-se o pacote LQMM – Linear Quantile Mixed Model (Geraci & Bottai, 2014). Nesse modelo as covariáveis podem exercer efeitos distintos nos diferentes quantis da distribuição, alinhando-se ao que é apresentado na regressão de quantílica padrão (Koenker & Bassett, 1978), e que o grau de heterogeneidade não observada também pode ser caracterizado com parâmetros de variância θ específicos.

A abordagem de estimativa baseia-se na relação existente entre a minimização da regressão dos Mínimos Desvios Absolutos (MDA) e a maximização de uma probabilidade de Laplace. A estimação dos parâmetros é realizada usando a função de distribuição Asymmetric Laplace (AL) que foi introduzida no contexto da regressão quantílica por Yu, Lu e Stander (2001) e por Koenker e Machado (1999). A função maximiza o log da verossimilhança de uma regressão de Laplace. Isso equivale à minimização da soma ponderada dos resíduos absolutos (Koenker e Bassett, 1978).

A demonstração matemática de suporte desse procedimento de estimação está além do escopo deste trabalho. Detalhes técnicos sobre o algoritmo de estimação ver Geraci e Bottai (2011).

O pacote *lqmm* tem sido utilizado em diferentes pesquisas nos últimos anos, Bellens, Kim & Van den Noortgate, Wim & Damme, Jan. (2019) estudaram a relação entre o livro didático adotado por professores e o desempenho dos estudantes em matemática na avaliação internacional TIMSS (Tendências em Matemática e estudo da ciência). A eficácia da redução da pobreza é avaliada por (Scott et al., 2020). Costanzo, 2015, investiga os efeitos de um programa de formação de professores italianos no desempenho dos estudantes de matemática. Solari et al., 2017 pesquisaram os efeitos de um programa de leitura destinado a estudantes com dificuldade na leitura no Texas. No Brasil, Barbeta et al, 2018 utilizam o pacote *lqmm* para investigar os fatores associados ao desempenho escolar de estudantes do 8º ano do ensino fundamental do estado do Pará

Associado ao modelo de regressão quantílica, será utilizada neste trabalho a abordagem hierárquica conforme apresenta Geraci e Bottai (2014):

Seja:

$$Y_{ij} = \beta_{0j}^{(\tau)} + \beta_1^{(\tau)} X_{1ij} + \beta_2^{(\tau)} X_{2ij} + \dots + \beta_L^{(\tau)} X_{Lij} + e_{ij}^{(\tau)} \quad (\text{nível 1})$$

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

onde:

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

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

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

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

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

$e_{ij}^{(\tau)}$ 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 σ e de assimetria igual a k ; e

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

Neste trabalho, a heterogeneidade das escolas será controlada atribuindo-se o efeito aleatório de $\beta_{0j}^{(\tau)}$, isto é, o termo aleatório permite controlar a heterogeneidade entre as unidades escolares nos diferentes quantis. As estimações do modelo podem ser calculadas pelo método da máxima verossimilhança. Os resultados foram obtidos a partir do pacote *lqmm*, desenvolvido por Geraci (2016), disponível no *software* livre R.

4. BASE DE DADOS

Os dados utilizados neste estudo são provenientes do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). As informações dos estudantes, incluindo a proficiência, são derivados do Sistema de Avaliação da Educação Básica (Saeb) 2017. Os dados de formação docente dos professores são retirados do Censo Escolar coletado anualmente.

Participaram do Saeb 2017 estudantes da rede pública das escolas com pelo menos dez alunos matriculados nas etapas do Ensino Fundamental nos anos iniciais (5º ano) e do Ensino Fundamental nos anos finais (9º ano). Além disso, participaram ainda estudantes concluintes da educação básica, isto é, as 3ª e 4ª séries do Ensino Médio.

O público-alvo da avaliação compreendeu 7.038.403 milhões de estudantes de 107.352 escolas, nos 5.570 municípios brasileiros. Na Tabela 1, nota-se que 77,56% dos alunos previstos participaram do Saeb, sendo que a região Nordeste registrou, em termos percentuais, a maior participação dos estudantes previstos, 79,50%.

Tabela 1 - Número de estudantes previstos no Saeb 2017

Região	Estudantes previstos	Estudantes que responderam ao teste	%
Centro-Oeste	549.707	431.664	78,53%
Nordeste	2.016.432	1.603.051	79,50%
Norte	791.424	607.812	76,80%
Sudeste	2.723.615	2.121.045	77,88%
Sul	957.225	695.641	72,67%
Brasil	7.038.403	5.459.213	77,56%

Fonte: Elaboração própria.

A Tabela 2 evidencia a participação dos estudantes com notas válidas por região e etapa avaliada. Observa-se um decréscimo no quantitativo de participantes do 5º para o 9º do Ensino Fundamental; e do Ensino Fundamental para o 3º ano do Ensino Médio. Esse decréscimo é reflexo das taxas de reprovação e abandono.

Tabela 2 – Número de estudantes que participaram do Saeb 2017 por etapa de ensino

Região	5º Ano Ensino Fundamental	9º Ano Ensino Fundamental	3ª Série Ensino Médio	TOTAL
Centro-Oeste	172.464	154.167	105.033	431.664
Nordeste	625.876	537.025	440.150	1.603.051
Norte	264.023	195.266	148.523	607.812
Sudeste	830.874	672.642	617.529	2.121.045
Sul	299.701	247.393	148.547	695.641
Brasil	2.192.938	1.806.493	1.459.782	5.459.213

Fonte: Elaboração própria.

Este trabalho concentra-se nos estudantes concluintes do Ensino Fundamental nos anos finais. Como a avaliação é realizada a cada dois anos, uma parte dos estudantes avaliados em 2017 participou da avaliação em 2013 quando cursava o 5º do Ensino Fundamental. Atendem a esse requisito os estudantes que não foram retidos ou não desistiram dos estudos no período 2017-2013. O desempenho anterior dos estudantes será incluído no modelo, para controlar o viés.

A Tabela 3 mostra o número de alunos com notas válidas em 2017 e 2013, sendo que a quarta coluna registra o número de alunos com notas válidas nas duas edições da avaliação.

Tabela 3 – Estudantes com notas válidas em 2017 e 2013

Região	Saeb 2017 – Estudantes COM NOTA VÁLIDA (9º ANO)	Saeb 2013 – Estudantes COM NOTA VÁLIDA (5º ANO)	Estudantes COM NOTAS VÁLIDAS nas duas edições
Centro-Oeste	154.167	172.711	89.217
Nordeste	537.025	557.085	230.365
Norte	195.266	248.733	105.230
Sudeste	672.642	798.952	431.776
Sul	247.393	294.185	146.134
Brasil	1.806.493	2.071.666	1.002.722

Fonte: Elaboração própria.

Após o tratamento da base de dados relacionada aos estudantes, e como o objetivo central deste trabalho é medir o efeito da formação dos professores no desempenho dos alunos, foram identificados todos os professores que lecionaram matemática no período de 2017-2013. O quadro a seguir mostra a relação entre professores, turma, escolas e alunos.

Tabela 4 – Estudantes, turmas, escolas e professores identificados no período de 2017-2013

Grupo	9º Ano	8º Ano	7º Ano	6º Ano
Estudantes	990.306	978.182	973.924	977.546
Turmas	74.173	84.979	92.185	96.448
Professores	47.137	54.432	59.427	57.315
Escolas	34.564	38.167	39.186	39.168

Fonte: Elaboração própria.

Nesta etapa, identificou-se casos de alunos que, durante o ano letivo, tiveram mais de um professor de matemática. Nesses casos, como não é possível identificar o tempo que cada professor lecionou para turma, esses registros foram excluídos da análise. A amostra final compreende os alunos que tiveram apenas um professor de matemática em cada etapa do Ensino Fundamental nos anos finais, isto é, do 6º ao 9º ano. Além disso, foram retirados da amostra os alunos da rede federal, devido ao número reduzido de estudantes matriculados nesse grupo.

A Tabela 5 compara o número de estudantes do 9º ano avaliados em 2017 com o número de alunos que compõe a amostra final utilizada neste estudo.

Tabela 5 – Estudantes que compõem a base de dados final

Região	Estudantes do 9º Ano Ensino Fundamental AVALIADOS	Estudantes do 9º Ano Ensino Fundamental AMOSTRA	%
Centro-Oeste	154.167	58.877	38,19%
Nordeste	537.025	182.423	33,97%
Norte	195.266	84.373	43,21%
Sudeste	672.642	349.429	51,95%
Sul	247.393	119.804	48,43%
Brasil	1.806.493	794.906	44,00%

Fonte: Elaboração própria.

Optou-se por incluir no modelo um número limitado de variáveis de controle, privilegiando aquelas com significância estatística comumente referenciadas na literatura de fatores associados ao desempenho escolar, como em Alves e Soares (2008, 2013); Laros et. al.

(2010). Sendo assim, a base de dados foi constituída com informações dos questionários contextuais aplicados aos alunos, diretores e escolas.

No estudo de fatores associados do Saeb (2016) disponível na Internet, são descritas várias medidas e indicadores construídos com base nos questionários contextuais, especialmente nos respondidos pelos estudantes.

Em estudos anteriores, com base na Teoria da Resposta ao Item (TRI), foram construídas várias medidas de escala gradual (Andrade; Tavares; Valle, 2000). Dentre essas medidas, três foram usadas neste estudo:

- nível socioeconômico da família do estudante (NSE);
- indicador de infraestrutura escolar; e
- indicador de estrutura pedagógica.

O indicador de nível socioeconômico contempla as variáveis de posse de bens e estrutura familiar. Nesse grupo, considerou-se a posse de TV em cores, rádio, geladeira, vídeo cassete ou DVD, freezer, banheiro, automóvel, máquina de lavar, escolaridade dos pais e a disponibilidade de empregada doméstica.

Foram incluídas no banco de dados variáveis que serão utilizadas como controle das características individuais dos estudantes, como: sexo, cor/raça, indicador de reprovação, indicador de gosto pela matemática, dentre outros. A relação completa das variáveis utilizadas para controle neste estudo podem ser encontradas no Quadro 1.

Em relação às variáveis de controle das características associadas às escolas, optou-se pelos indicadores desenvolvidos por pesquisadores da Universidade Federal de Minas Gerais, divulgados na publicação “Qualidade da Infraestrutura das Escolas públicas do Ensino Fundamental no Brasil” (2019), que contemplam os diferentes aspectos do ambiente escolar, inclusive levando em conta as diferenças regionais e territoriais que afetam as escolas brasileiras. O primeiro indicador é o de infraestrutura geral e agrega outros 11 (onze) indicadores: serviços básicos, instalações do prédio, prevenção de danos, conservação, conforto, ambiente prazeroso, espaços pedagógicos, equipamentos para apoio administrativo, equipamentos para apoio pedagógico, acessibilidade, ambiente para atendimento educacional especializado. O segundo indicador escolhido refere-se ao conjunto de informações sobre os equipamentos de apoio pedagógicos disponíveis nas unidades escolares. Informações adicionais sobre a metodologia adotada na elaboração desses indicadores, bem como o detalhamento das variáveis que os compõem podem ser vistos em Alves e Xavier (2018).

A variável que indica a formação adequada em matemática foi construída a partir dos critérios estabelecidos pelo Inep, conforme consta na Nota Técnica Inep/Deed nº 020/2014, de 21 de novembro de 2014. Sendo assim, a formação será considerada adequada quando o professor de matemática possuir formação superior de licenciatura em matemática, ou bacharelado em matemática com curso de complementação pedagógica concluído.

A Tabela 6 revela que no período de 2014 a 2017, considerando os estudantes que compõem a amostra final deste estudo, houve um incremento dos professores com a formação adequada em matemática de, aproximadamente, 10 pontos percentuais. Contudo, ainda nesse quesito, persistem as desigualdades regionais. A região Nordeste registra o menor percentual de estudantes que tiveram professores de matemática com formação adequada na disciplina. Por outro lado, é a região que apresentou a maior evolução no período, passando de 42,13%, em 2014, para 60,45% em 2017, um aumento de 18,32 p.p. no período.

Tabela 6 – Professores com formação adequada em matemática

Região	2017 (9º Ano)	2016 (8º Ano)	2015 (7º Ano)	2014 (6º Ano)
Centro-Oeste	72,58%	72,32%	70,47%	65,81%
Nordeste	60,45%	56,15%	49,64%	42,13%
Norte	81,61%	78,33%	74,37%	67,90%
Sudeste	80,16%	77,73%	76,44%	72,93%
Sul	74,91%	72,74%	71,94%	68,54%
Brasil	74,44%	71,69%	68,95%	64,14%

Fonte: Elaboração própria.

Como mencionado anteriormente, a base de dados não contempla todos os alunos da rede pública matriculados no 9º ano do ensino fundamental. Estão fora da amostra os estudantes que não realizaram as provas nas duas edições (2013 – 5º ano e 2017 – 9º), seja por ausência nos dias de realização dos testes, por abandono ou por reprovação. Além disso, foram excluídos os registros dos estudantes que tiveram mais de um professor de matemática em 2017. Após esse tratamento, é possível comparar as características dos estudantes da amostra final, que compreende estudantes com nota válida nas duas edições (2017 e 2013), com o universo de estudantes do 9º ano do ensino fundamental da rede pública que participaram da Avaliação em 2017.

A tabela 7 apresenta as características dos estudantes e escolas e, como esperado, há diferença em termos das médias no perfil dos dois grupos, estatisticamente significantes a 1%.

Essa diferença compreende quase todas as características analisadas, resultando em estimativas superestimadas, especialmente para a situação socioeconômica, sendo a característica mais afetada. Esse resultado não surpreende, pois sabe-se que os estudantes mais pobres são também os mais afetados pela reprovação. Esse é o principal motivo de perdas de registros na amostra final em relação à base de dados original. Em relação às características de formação dos professores, foco de análise desse estudo, as diferenças apuradas, embora estatisticamente significativas, são pouco acentuadas, o que não compromete os resultados esperados para essas variáveis

Tabela 7 – Características dos estudantes, professores e escolas

	Estudantes COM NOTA VÁLIDA (9º ANO) 1.806.493	Estudantes do 9º Ano Ensino Fundamental AMOSTRA FINAL 794.906	Teste de Igualdade de Médias
	Média	Média	
Localização da Escola			
Região N	0,1087	0,1061	
Região NE	0,2980	0,2295	***
Região SE	0,3847	0,4396	***
Região S	0,1362	0,1507	***
Região CO	0,0723	0,0741	***
Zona urbana	0,8844	0,9194	***
In Capital	0,1671	0,1474	***
Rede			
Municipal	0,4481	0,4433	***
Estadual	0,5396	0,5567	***
Infraestrutura escolar			
Equi_ap_pedag10	5,7634	5,9307	***
Infrageral10	6,7932	6,8625	***
Características do estudante			
Sexo fem.	0,5110	0,5357	***
Cor Preta	0,1181	0,0327	***
NSE	0,0357	0,1125	***
Reprovou	0,2943	0,1320	***
Gosta Mat	0,6533	0,6695	***
Características do professor de Matemática (9º ano)			
Possui Especialização	0,4893	0,4887	***
Formação em matemática	0,7290	0,7300	***

Fonte: Elaboração própria. Estimativas estatisticamente significativas a 1 %.

5. RESULTADOS E DISCUSSÃO

Nessa seção, são apresentados os resultados de regressões quantílicas hierárquicas aplicadas aos estudantes de matemática do 9º ano do Ensino Fundamental do Saeb 2017. As estimativas foram geradas por região e correspondem ao efeito da formação adequada do professor nos diferentes quantis e percentis do desempenho dos estudantes.

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. O método de estimação foi o de máxima verossimilhança restrita, usando o pacote lme4 (Bates et al., 2015). A inserção de variáveis no modelo teve como referência o proposto por Hox (2010):

- a) inserção do modelo nulo que serve para a comparação com os modelos subsequentes e para o cálculo da correlação intraclasse;
- b) inserção no modelo das variáveis explicativas fixas no nível do estudante. Primeiro, introduziu-se a variável de interesse, se o estudante teve professor com formação adequada, e depois, a variável que agrega a proficiência anterior do estudante;
- c) inserção no modelo das demais variáveis de suporte familiar; e
- d) no modelo final, a inserção de variáveis explicativas no nível da escola.

A análise do modelo nulo (Modelo 0) é importante para se ter uma estimação de linha de base da variância (Tabela 8). Dada a variância inicial, é possível calcular o percentual de variância explicada com a inclusão no modelo das variáveis explicativas.

A correlação intraclasse ρ indica a proporção da variância explicada no nível do grupo em comparação com a variância total e pode ser obtida pela equação:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2}$$

Nessa equação, σ_{u0}^2 é a variância dos resíduos u_{0j} do nível mais elevado (nível 2 - da escola) e σ_e^2 é a variância dos resíduos e_{ij} do nível mais baixo (nível 1 - do estudante). É com base nesse modelo que se calcula a correlação intraclasse (Intraclass Correlation

Coefficient – ICC), que varia entre 0 e 1 e indica a necessidade ou não de utilizar um modelo multinível.

Tabela 8 – Modelo apenas com o intercepto (modelo nulo)

Modelo	M1: Apenas Intercepto	
Efeito fixo	Coefficiente (EP)	Estat-t
Intercepto	254,37 (0,13)	1.848,28
Efeito aleatório		
Efeito randômico - nível 1 - σ_e^2	1.842,1 (42,92)	
Variância – intercepto - nível 2 - σ_{u0}^2	446,6 (21,13)	
Correlação intraclasse	0,19	

Fonte: Elaboração própria. Erro padrão entre parênteses. Todas as estimativas são estatisticamente significativas a 5%.

No modelo nulo em questão, verificou-se um intercepto de 254,37, o qual representa o valor médio da proficiência em matemática dos alunos do 9º ano do ensino fundamental em 2009. A correlação intraclasse é de 0,19, o que significa que aproximadamente 19% da variância nas proficiências dos estudantes em matemática pode ser atribuída ao nível da escola. O valor relativamente elevado da correlação intraclasse justifica a utilização de uma abordagem multinível de análise em vez de uma regressão tradicional. Há autores que consideram que, quando esse valor é superior a 10%, justifica-se a utilização do modelo multinível (Ferrão, 2003; Lee, 2008). Todavia, dependendo das circunstâncias, mesmo valores de ICC entre 5% e 10% já são indicativos de uma dependência entre níveis (Bickel, 2007).

Em seguida, a variável de interesse foi adicionada ao modelo. Essa variável *dummy* indica se o estudante teve professor com a formação adequada em matemática em todas as etapas do ensino fundamental. Com a inserção da variável, obtém-se o modelo M2, resumido na Tabela 9.

Tabela 9 – Modelo com o intercepto e a adequação da formação do professor

Modelo	M2: Intercepto e variável de interesse	
Efeito fixo	Coefficiente (EP)	Estat-t
Intercepto	253,36 (0,14)	1.732,63
Formação adequada do professor	2,97 (0,14)	19,86
Efeito aleatório		
Efeito randômico - nível 1 - σ_e^2	1.841,90 (42,92)	
Variância – intercepto – nível 2 - σ_{u0}^2	441, 10 (21,00)	
Correlação intraclasses	0,19	

Fonte: Elaboração própria. Erro padrão entre parênteses. Todas as estimativas são estatisticamente significativas a 5%.

A inclusão dessa variável provoca poucos efeitos nas estimativas. Entretanto, o modelo indica que há uma relação positiva e estatisticamente significativa entre a formação adequada do professor de matemática sobre o aprendizado dos estudantes na mesma disciplina. A diferença no desempenho médio dos estudantes cujos os professores tinham formação adequada é de 2,97 pontos na proficiência dos estudantes.

A proficiência do estudante em matemática no 5º ano do Ensino Fundamental foi adicionada como variável explicativa no modelo M3 (Tabela 10). Isso é importante pois a ausência desse controle pode gerar resultados imprecisos. Ademais, algumas características não observáveis dos estudantes, como habilidade inata e motivação, tendem a sofrer pouca variação ao longo do tempo, sendo que os efeitos das variáveis omitidas serão capturados pela inclusão da proficiência anterior do estudante.

Cabe ressaltar que a inclusão dessa variável de controle altera a interpretação dos resultados. Ao avaliar o desempenho dos estudantes do 9º ano, depois de levar em consideração a variação associada à proficiência no 5º ano, o modelo passa a avaliar efetivamente o progresso entre o 5º e o 9º ano. Essa abordagem pode gerar interpretações mais confiáveis, pois permite ao pesquisador analisar o incremento da variável de interesse sobre a proficiência adquirida pelo estudante no período.

Tabela 10 – Modelo com o intercepto, proficiência anterior e adequação da formação do professor

Modelo	M3: Com variável proficiência anterior	
Parte fixa	Coefficiente (EP)	Estat-t
Intercepto	123,62 (0,21)	591,13
Formação adequada do professor	0,95 (0,11)	8,46
Proficiência em matemática no 5º ano	0,60 (0,00)	697,22
Parte aleatória		
Efeito randômico - nível 1 - σ_e^2	1.168,06 (34,18)	
Variância – intercepto – nível 2 - σ_{u0}^2	149,80 (12,24)	
Correlação intraclasse	0,11	

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses. Todas as estimativas são estatisticamente significativas a 5%.

Nesse modelo, a proficiência anterior reduz a variação do erro residual no primeiro nível (nível estudante) para 1.168,1 e na variação do segundo nível (escola) para 149,9. Essa diferença pode ser interpretada como a variação explicada pela inclusão da variável “proficiência anterior” do estudante ao modelo M3. Para calcular uma estatística análoga ao R^2 múltiplo, devemos expressar essa diferença como uma proporção da variação total do erro.

A análise será mais adequada quando realizada separadamente, isto é, no nível do aluno e no nível da escola, tendo como referência os modelos M1 e M3. Para a proporção de variância explicada no nível do aluno tem-se (Raudenbush e Bryk, 2002):

$$R_1^2 = \frac{\sigma_{e\backslash b}^2 - \sigma_{e\backslash m}^2}{\sigma_{e\backslash b}^2}$$

Onde $\sigma_{e\backslash b}^2$ é a variância residual do nível de aluno do modelo nulo (M1), e $\sigma_{e\backslash m}^2$ é a variância residual do modelo de comparação (M3), também no nível do aluno. Assim:

$$R_1^2 = \frac{1.842,1 - 1.168,1}{1.842,1} = 0,37$$

Em relação a proporção de variância explicada no nível da escola (Raudenbush e Bryk, 2002), tem-se que:

$$R_2^2 = \frac{\sigma_{u0|b}^2 - \sigma_{u0|m}^2}{\sigma_{u0|b}^2}$$

Onde $\sigma_{u0|b}^2$ é a variância residual do nível da escola do modelo nulo (M1), e $\sigma_{u0|m}^2$ é a variância residual do modelo de comparação (M3), no nível da escola. Assim:

$$R_2^2 = \frac{446,6 - 150,1}{446,6} = 0,66$$

A variância explicada no modelo (M3) é de 37% no nível do aluno e de 66% no nível de escola. Os valores encontrados para o modelo (M4), que leva em conta, além das variáveis indicadas no modelo (M3), as variáveis de controle indicadas no Quadro 1, correspondem a 40% para o nível de aluno, e 71% para o nível da escola. Isso indica que o modelo (M4), depois de incluídas as variáveis de controle, se ajusta melhor aos dados, em comparação aos modelos iniciais.

No modelo subsequente (M4), foram incluídas as variáveis de controle nos dois níveis, isto é, estudante e escola, conforme Quadro 1. Optou-se pela utilização das variáveis não padronizadas, a fim preservar as interpretações das escalas originais do Saeb.

Quadro 1 – Descrição das variáveis utilizadas no modelo

Variáveis	Descrição
Proficiência Mat	Proficiência em matemática no 9º Ano em 2017
Formação Adequada	Dummy igual a 1 se os professores que lecionaram matemática durante todo o EF possuíam formação adequada em Matemática, e 0 caso contrário
Pós-Graduação	Dummy igual a 1 se os professores que lecionaram matemática durante todo o EF possuíam curso de pós-graduação em Matemática, e 0 caso contrário
Proficiência Mat – anterior	Proficiência em matemática no 5º Ano em 2013
Sexo_feminino	Dummy igual a 1 se o estudante é do sexo feminino, e 0 caso contrário

continua

Variáveis	Descrição
Cor_preta	Dummy igual a 1 se o cor declarada pelo estudante é preta, e 0 caso contrário
Reprovou	Dummy igual a 1 e estudante já reprovou, e 0 caso contrário
Gosta Mat	Dummy igual a 1 se o estudante gosta de matemática, e 0 caso contrário
NSE	Indicador de nível socioeconômico do aluno.
Dep_estadual	Dummy igual a 1 se a dependência administrativa da escola é estadual, e 0 caso contrário (escola)
Zona_urbana	Dummy igual a 1 se a localização da escola é na zona urbana, e 0 caso contrário (escola)
IN_Capital	Dummy igual a 1 se a localização da escola é na capital, e 0 caso contrário (escola)
Infrageral10	Indicador de infraestrutura geral (escola)
Equi_ap_pedag10	Indicador de equipamento para apoio pedagógico disponíveis na escola (escola)

Fonte: Elaboração própria.

Os resultados serão apresentados a seguir em nível nacional, evidenciando os efeitos nas redes de ensino estadual e municipal, além disso, serão apresentados quadro específicos para cada região em anexo.

Tabela 11 – Modelo Brasil

Modelo (M4)	Estadual e Municipal			Estadual			Municipal		
	Coeficiente	EP	estat-t	Coeficiente	EP	estat-t	Coeficiente	EP	estat-t
Parte fixa									
Intercepto	123,91	0,77	160,70	125,42	1,12	112,37	118,47	1,16	101,93
Formação Adequada	0,56	0,11	5,00	0,51	0,14	3,61	0,56	0,18	3,11
Pós-Graduação	1,80	0,16	11,50	2,37	0,20	11,65	1,19	0,24	4,95
Proficiência Mat - anterior	0,55	0,00	611,93	0,56	0,00	469,08	0,54	0,00	392,18
Sexo_feminino	-6,51	0,08	-82,70	-5,58	0,10	-53,16	-7,71	0,12	-64,81
Cor_preta	-3,61	0,22	-16,16	-3,97	0,30	-13,08	-3,18	0,33	-9,67
Reprovou	-12,64	0,12	-103,96	-12,75	0,17	-74,47	-12,58	0,17	-72,81
Gosta Mat	12,72	0,08	150,15	12,80	0,11	114,76	12,58	0,13	96,77
Nse	2,64	0,06	46,32	2,94	0,08	38,58	2,27	0,09	26,53
Dep_estadual	-4,19	0,17	-24,33						
Zona_urbana	-1,43	0,26	-5,55	-1,12	0,50	-2,24	-1,90	0,32	-5,95

continua

continuação

Modelo (M4)	Estadual e Municipal			Estadual			Municipal		
	Coefficiente	EP	estat-t	Coefficiente	EP	estat-t	Coefficiente	EP	estat-t
Parte fixa									
IN_Capital	-2,78	0,25	-11,36	-2,99	0,29	-10,38	-3,10	0,44	-7,10
Infrageral	1,48	0,13	11,36	0,45	0,17	2,69	2,68	0,20	13,18
Equi_ap_pedag	0,25	0,05	5,05	0,13	0,06	1,95	0,35	0,08	4,52
Parte aleatória									
	Estadual e Municipal			Estadual			Municipal		
Efeito Randômico - nível $1 - \sigma_e^2$	1.090,3 (33,02)			1.089,3 (33,00)			1.090,3 (33,02)		
Variância – intercepto - nível 2 - σ_{u0}^2	128,9 (11,35)			102,4 (10,12)			157,7 (12,56)		
Correlação intraclasse	0,105			0,086			0,126		

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses.

A Tabela 11 apresenta os resultados do modelo para todo o universo do banco de dados disponível. Os resultados desse modelo mostram que a formação do professor melhora o desempenho no teste de matemática nas duas redes de ensino, sendo que a formação adequada tem efeito um pouco maior na rede municipal, enquanto a especialização do professor (curso de pós-graduação) apresenta resultado mais acentuado na rede estadual. Ainda sobre esse quesito, quando comparado os efeitos dos professores com formação adequada, daqueles que possuem especialização, observa-se que a segunda variável exerce maior influência sobre o desempenho dos estudantes; sendo que na rede estadual esse efeito é, aproximadamente, quatro vezes maior do que os efeitos atribuídos à primeira variável explicativa.

Os resultados por região, apresentados no Anexo, evidenciam que os efeitos da formação do professor nas regiões Sul e Sudeste são similares ao apresentado em nível nacional, em que os efeitos são estatisticamente significativos em todos os grupos pesquisados, sendo mais forte na rede municipal. Nas regiões Norte e Centro-Oeste, as variáveis de interesse mostram-se significante apenas na rede municipal. Por outro lado, na região Nordeste há efeitos positivos, entretanto os resultados não são significativos nos grupos analisados.

Em relação aos resultados obtidos das variáveis utilizadas como controle, pode-se afirmar que há semelhança com os resultados de outros estudos que versam sobre os fatores que afetam o desempenho dos estudantes, como em Laros (2010) e Alves e Soares (2008). Cabe mencionar os resultados encontrados para variável de gênero: todos os grupos apresentam efeitos negativos para o sexo feminino. Na região Nordeste, em média as meninas tiveram 9,03

pontos a menos comparados aos meninos. Esse valor é quase o dobro do observado na região Sudeste, que registra para as meninas 4,58 pontos a menos que os meninos.

Em relação a cor/raça, na região Sul os estudantes que se declaram pretos obtiveram 4,90 pontos a menos comparados aos demais grupos. É a região que apresenta o maior efeito para essa variável. Na região Nordeste, por exemplo, os estudantes de cor preta registram em média, 2,17 pontos a menos nas provas de matemática do Saeb.

O coeficiente de correlação intraclasse (ICC) também é mais forte na região Nordeste: aproximadamente 14,2% da variância nas proficiências dos estudantes em matemática pode ser atribuída ao nível da escola, sendo que na rede municipal esse percentual é de 15,3%. Nas demais regiões, o coeficiente é de aproximadamente 8%.

Os efeitos são observados no nível médio de proficiência, mas esses resultados não respondem como a formação do professor pode influenciar os estudantes de baixa ou alta proficiência. Será que os efeitos são os mesmos nas diferentes regiões do país?

De um modo geral, o que se busca responder neste estudo é se o efeito da formação do professor é maior em estudantes de alto desempenho ou de baixo desempenho. Nesse sentido, a análise a seguir permite que os coeficientes sejam diferentes para diversos quantis condicionais da distribuição das proficiências visando responder a questão levantada no parágrafo anterior.

Na Tabela 12, estão apresentados vários quantis incondicionais com os resultados do Saeb (2017) para a prova de matemática. O quantil 25%, igual a 240,61, indica que 25% dos estudantes têm proficiência em matemática igual ou inferior a 240,61. O quantil 75%, por sua vez, mostra que 75% dos estudantes têm proficiência em matemática igual ou inferior a 300,53 pontos na escala do Saeb. Os quintis 10% e 90% separam, respectivamente, os 10% de menor proficiência e os 10% de maior proficiência.

Tabela 12 – Proficiência dos estudantes em matemática do 9º ano do Ensino Fundamental no Saeb (2017)

Região	QUANTIL				
	10%	25%	50%	75%	90%
Sul	212,71	240,61	270,86	300,53	325,85
Sudeste	199,30	229,88	262,84	295,91	324,09
Centro-Oeste	201,63	229,87	260,79	292,63	319,68
Nordeste	187,17	214,20	244,32	277,82	309,45
Norte	188,42	214,51	243,50	275,01	304,55
Brasil	196,16	225,30	257,55	291,13	319,88

Fonte: Elaboração própria.

Os resultados das estimativas pela regressão quantílica são apresentados na tabela 13. As variáveis de interesse são as *dummies* “Formação Adequada” e “Pós-Graduação” que recebem valor 1 quando o estudante, durante toda a trajetória acadêmica do ensino fundamental nos anos finais da educação básica, teve professor de matemática com formação adequada na área e professor com curso de pós-graduação. As variáveis de controle utilizados no modelo nos níveis de escola e alunos estão relacionadas no quadro 1 e são os mesmos apresentados na tabela 11. Os resultados por região e por rede de ensino estão apresentados no Anexo.

Tabela 13 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, no Brasil.

Escolas das redes estadual e municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,77 *	0,54 *	0,37 *	0,36 .	0,36
	(0,37)	(0,24)	(0,15)	(0,20)	(0,32)
Pós-Graduação	1,82 **	2,28 ***	2,14 ***	2,11 ***	1,20 **
	(0,58)	(0,28)	(0,19)	(0,31)	(0,41)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede estadual					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,98*	0,51*	0,37*	0,51*	0,99**
	(0,41)	(0,25)	(0,16)	(0,22)	(0,36)
Pós-Graduação	2,70***	2,62***	2,29***	2,21***	1,94***
	(0,52)	(0,29)	(0,31)	(0,34)	(0,38)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	1,30*	0,29	0,44*	0,23	0,22
	(0,55)	(0,30)	(0,18)	(0,28)	(0,47)
Pós-Graduação	2,20**	1,34**	1,15**	1,38**	0,85
	(0,65)	(0,48)	(0,30)	(0,45)	(0,56)
Controles: Alunos e escolas, conforme quadro 1.					

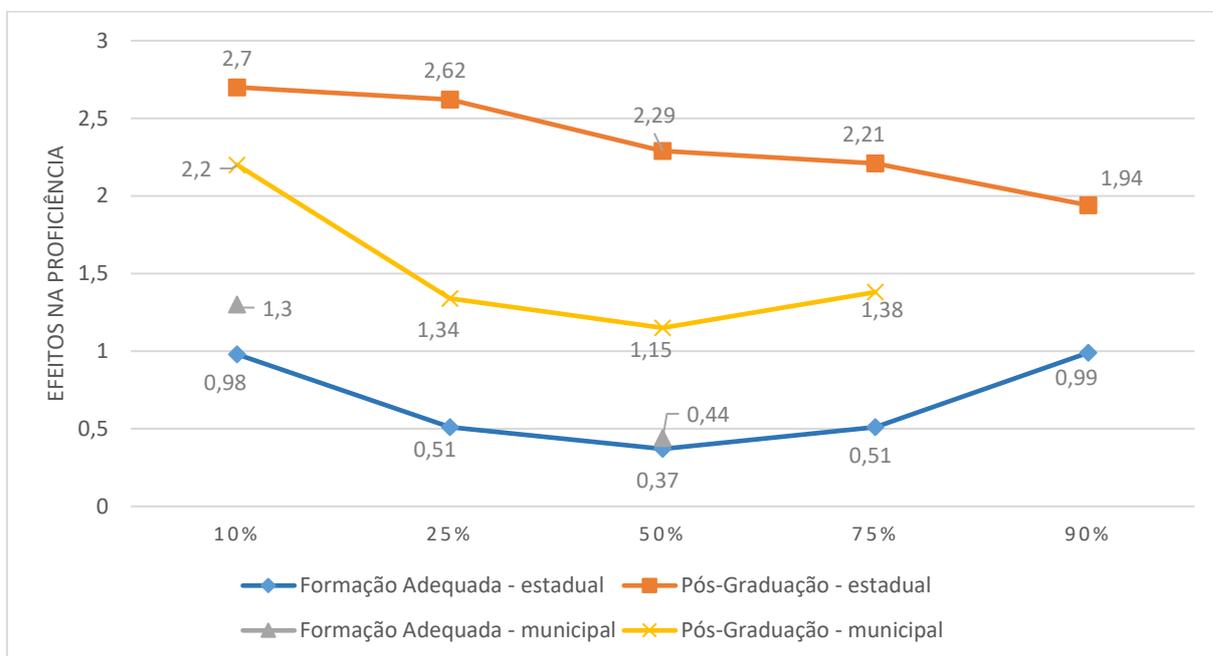
Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses. Significância: 0 ‘****’ 0.001 ‘***’ 0.01 ‘**’ 0.05 ‘.’ 0.1 ‘.’ 1.

Os resultados indicam efeitos positivos e estaticamente significativos das variáveis estudadas. Assim como na rede municipal, os estudantes atendidos pela rede estadual de mais baixa proficiência, que estão no 10% quantil, são mais afetados pela formação adequada do professor; sendo que na rede municipal esse efeito é superior ao registrado na rede sob gestão do governo estadual. Observa-se que a rede estadual também apresenta efeito positivo no outro extremo, isto é, no grupo de estudantes de maior proficiência, 90% quantil. Em todos os pontos da distribuição de proficiência, essa variável apresenta maior efeito na rede municipal quando comparada à estadual.

Quando são analisados os efeitos no desempenho dos estudantes cujos os professores possuem pós-graduação, observa-se, novamente, que os estudantes de mais baixa proficiência são os mais afetados por essa variável. Na rede estadual no 10% quantil o efeito é de 2,7 pontos e de 1,94 no 90% quantil; e na rede municipal de 2,2 pontos no 10% quantil e 1,38 pontos no quantil 75%. Os resultados podem ser vistos no gráfico 1.

Gráfico 1 – Efeitos da formação adequada dos professores e do curso de pós-graduação na proficiência dos estudantes

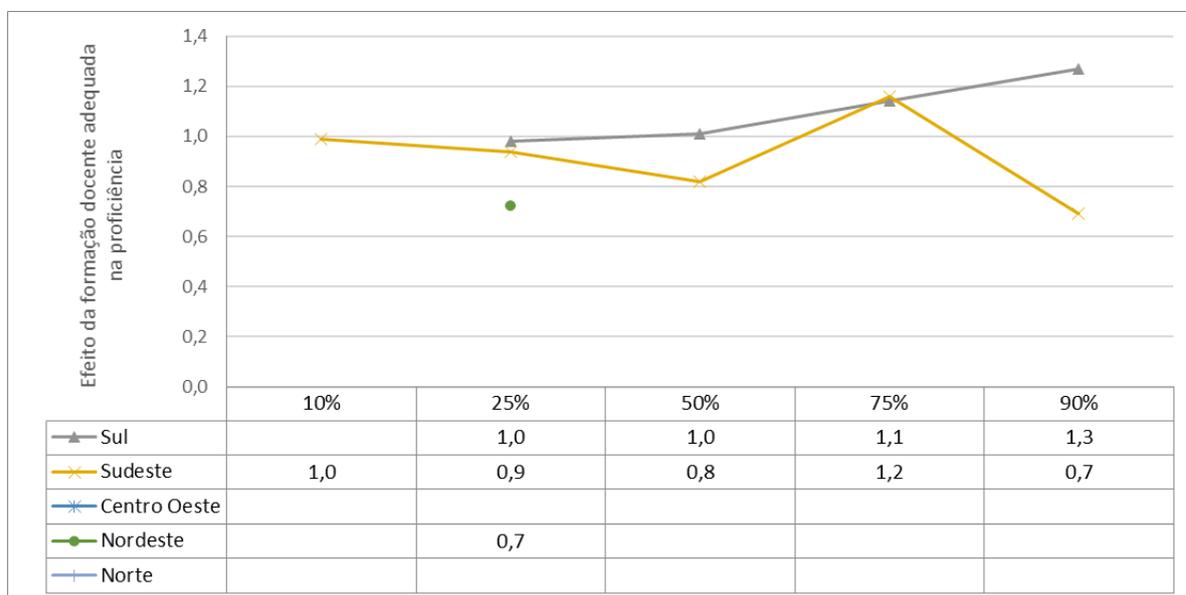


Fonte: Elaboração própria.

No Gráfico 2, são apresentadas apenas as estimativas estatisticamente significantes ao nível de 0,05. Os efeitos da variável “formação adequada” são positivos e estatisticamente

significativos em todos os percentis estudados da região Sudeste, sendo ligeiramente mais forte no quantil 75%, que separa os 25% dos estudantes com maior proficiência. O efeito estimado nesse grupo foi de 1,2 pontos. Na região Sul, os efeitos da formação adequada do professor afetam mais os estudantes de maior proficiência, observa que o aumento de um desvio da variável leva a um aumento de 1 ponto na proficiência mediana, e de 1,3 no quantil 90%, que separa os 10% dos estudantes de maior proficiência na região.

Gráfico 2 – Efeitos da formação docente adequada na proficiência dos estudantes do 9º ano



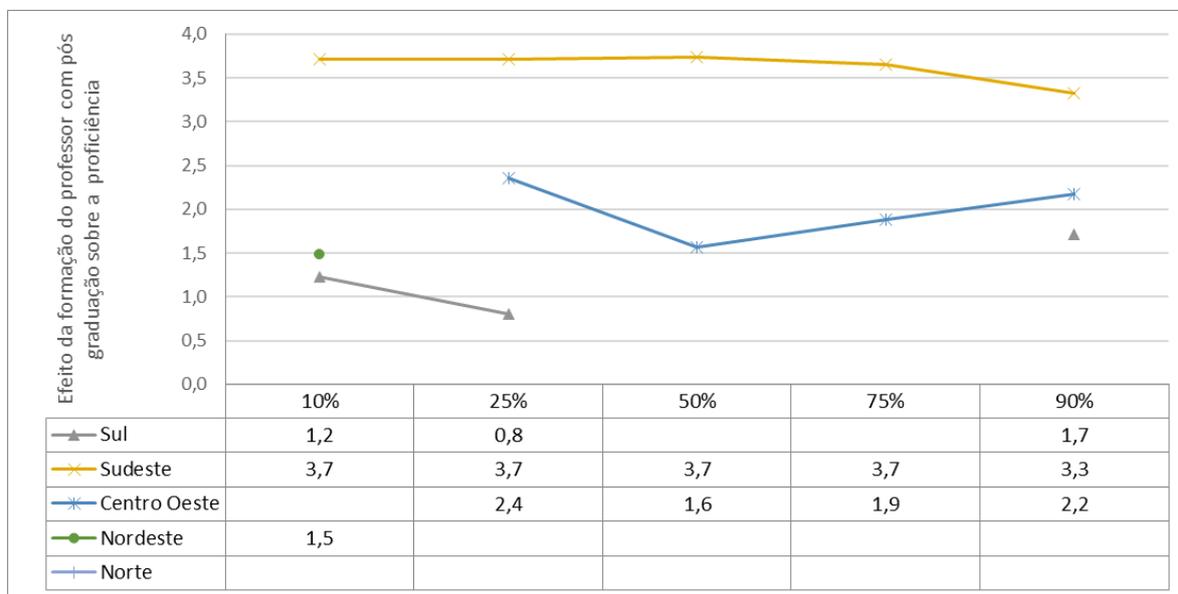
Fonte: Elaboração própria.

Na região Nordeste, registrou-se efeito contrário do ocorrido na região Sudeste. Entre os estudantes nordestinos os efeitos da formação do professor de matemáticas foram mais fortes no quantil inferior, ou seja, aqueles de menor proficiência. Nessa região, os efeitos da formação adequada dos professores que lecionaram matemática é de aproximadamente 1 ponto, sendo registrado apenas no quantil 25%. Nas regiões Norte e Centro-Oeste, os resultados encontrados não são estatisticamente significativos.

No Gráfico 3, são apresentados os efeitos no desempenho dos estudantes cujos os professores possuem curso de pós-graduação. Apenas na região Norte não foram encontrados efeitos dessa variável. Na região Centro-Oeste, os efeitos da formação do professor com pós-graduação afetam mais os estudantes de maior proficiência, observa que o aumento de um

desvio da variável leva a um aumento de 1,6 pontos na proficiência mediana, de 1,9 no quantil 75%, e de 2,2 pontos no 90% de maior proficiência na região.

Gráfico 3 – Efeitos da formação com curso de pós-graduação na proficiência



Fonte: Elaboração própria.

Na região Nordeste observa-se que os estudantes de mais baixa proficiência são afetados pela variável no 10% quantil, o efeito nesse grupo é de 1,5 pontos.

Assim como a variável anterior, os efeitos da formação do professor com pós-graduação, na região Sudeste, afetam todos os quantis pesquisados, sendo superior a 3 pontos em todos eles. É a região que apresenta o maior efeito para essa variável.

Os resultados para a região Sul mostram que efeitos estão concentrados nos estudantes dos quantis extremos, 10%, 25% e 90%, sendo mais forte para os estudantes do quantil 90, sendo que o efeito estimado nesse grupo foi de 1,7 pontos na região.

No entanto, uma limitação dos resultados apresentados anteriormente é a possibilidade de endogenidade nos modelos, produzindo estimativas inconsistentes devido a não aleatorização dos estudantes nas turmas e/ou na designação dos respectivos professores.

As escolas podem adotar diferentes estratégias na composição das turmas, utilizando informações das características dos estudantes (idade, reprovação, fraco ou alto desempenho). O mesmo pode ocorrer na escolha do professor de cada turma. Os professores mais antigos podem eventualmente ter o direito de escolher a turma que lecionará, dando preferência a turmas de estudante de alto desempenho, ou outro critério que julgar mais adequado para o

desempenho de suas funções. A gestão escolar também pode interferir de forma não aleatória na atribuição das turmas aos professores, utilizando-se de informações dos estudantes e dos professores.

Embora alguns pesquisadores argumentem que a inclusão no modelo de variáveis de controles (por exemplo, desempenho do estudante em testes anteriores, características demográficas) seja suficiente para estimar o efeito do professor sobre o desempenho dos estudantes (Chetty, Friedman e Rockoff 2014a; Kane et al.2013), a designação aleatória de professores às turmas oferece uma alternativa de validar essa suposição.

Afim de testar a robustez do modelo e controlar a possível endogeneidade entre professores e alunos, estimados modelos para um grupo de escolas que possivelmente distribuíram os estudantes e professores na turmas de forma não homogênea. Esse grupo de escolas pode ser identificado com a seleção de registros da base de dados do questionário do Saeb, respondido pelos diretores em relação aos critérios adotados pela escola na organização das turmas (Questão 39) e na atribuição dos professores em cada turma (Questão 40). Conforme Quadro 2, foram selecionados os registros das observações destacadas em negrito.

Quadro 2 – Questões 39 e 40 do Questionário do Diretor – Saeb 2017

QUESTÃO - 39. NESTE ANO, QUAL FOI O PRINCIPAL CRITÉRIO UTILIZADO PARA A FORMAÇÃO DAS TURMAS NESTA ESCOLA?	QUESTÃO - 40. NESTE ANO, QUAL FOI O PRINCIPAL CRITÉRIO PARA A ATRIBUIÇÃO DAS TURMAS AOS PROFESSORES?
<p>(A) Homogeneidade quanto à idade (alunos com a mesma idade).</p> <p>(B) Homogeneidade quanto ao rendimento escolar (alunos com rendimento similar).</p> <p>(C) Heterogeneidade quanto à idade (alunos com idades diferentes).</p> <p>(D) Heterogeneidade quanto ao rendimento escolar (alunos com rendimentos diferentes).</p> <p>(E) Outro critério.</p> <p>(F) Não houve critério.</p>	<p>(A) Preferência dos professores.</p> <p>(B) Escolha dos professores, de acordo com a pontuação por tempo de serviço e formação.</p> <p>(C) Professores experientes com turmas de aprendizagem mais rápida.</p> <p>(D) Professores experientes com turmas de aprendizagem mais lenta.</p> <p>(E) Manutenção do professor com a mesma turma.</p> <p>(F) Revezamento dos professores entre as séries.</p> <p>(G) Sorteio das turmas entre os professores</p> <p>(H) Atribuição pela direção da escola.</p> <p>(I) Outro critério.</p> <p>(J) Não houve critério.</p>

Fonte: Questionário do diretor – Saeb 2017.

O primeiro modelo, cujos resultados são apresentados na Tabela 14, refere-se às escolas que não utilizaram as informações de idade e desempenho dos estudantes na formação das turmas homogêneas, supondo-se neste caso, que a distribuição tenha sido realizada de forma heterogênea ou aleatória.

O modelo seguinte, além de considerar apenas as escolas que supostamente organizam as turmas de forma heterogênea ou aleatória, também restringiu-se às escolas que designaram aleatoriamente os professores às turmas.

A magnitude dos efeitos apresentados na Tabela 14, é semelhante ou maior do que os resultados discutidos na Tabela 13. Deve-se destacar que para o grupo de escolas que não adota critério de homogeneidade na formação das turmas observou-se efeitos da formação adequada dos professores apenas na proficiência mediana dos estudantes, sendo que os demais quantis não apresentaram efeitos estatisticamente significantes. Os efeitos da variável de formação do professor com pós-graduação para esse grupo, são positivos e estatisticamente significativo, afetando mais fortemente os estudantes de mais baixa proficiência, que estão no 10% quantil.

O modelo mais restritivo em que, concomitante professores e estudantes são atribuídos de forma heterogênea ou aleatória, nos quantis com estimativas estatisticamente significantes, os efeitos são positivos. Destaca-se que os efeitos das duas variáveis de interesse sobre o desempenho dos estudantes nesse grupo mostraram-se mais acentuados quando comparados ao público indicado na Tabela 13. Isso sugere que o modelo inicial pode ter subestimado os efeitos das variáveis sobre o desempenho dos estudantes. Por outro lado, deve-se ressaltar que na amostra restrita, muitos registros foram perdidos, após os filtros mencionados anteriormente, que podem ter causado algum tipo de viés nos resultados desse grupo.

Tabela 14 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, no Brasil.

Escolas com organização das turmas não homogêneas – Questão 39					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,98 (1,13)	0,03 (0,86)	1,22* (0,62)	-0,04 (0,70)	1,36 (1,15)
Pós-Graduação	2,77* (1,62)	0,81 (1,19)	1,74 (1,27)	2,09* (0,93)	0,32 (1,61)
Controles: Alunos e escolas, conforme quadro 1.					

Continua

Escolas com organização das turmas não homogêneas e atribuição aleatória dos professores					
Questões 39 e 40					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	1,71 (1,56)	1,27 (1,18)	1,67* (0,87)	-0,28 (0,87)	0,00 (1,39)
Pós-Graduação	3,80* (1,90)	3,11* (1,60)	2,07 (1,61)	2,90* (1,70)	1,20 (2,20)
Controles: Alunos e escolas, conforme quadro 1.					

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses. Significância: 0 ‘*****’ 0.001 ‘***’ 0.01 ‘**’ 0.05 ‘*’ 0.1 ‘.’ 1.

6. CONSIDERAÇÕES FINAIS

Nos últimos anos, pesquisadores realizaram diversos estudos a fim de identificar intervenções inovadoras que auxiliassem as escolas na obtenção de melhores níveis de qualidade da educação.

Nesse mesmo período, o Brasil implementou uma moderna metodologia de avaliação em larga escala, que fornece informações necessárias para identificar o nível de eficiência das políticas públicas educacionais estabelecidas pelo Governo Federal. Além disso, o Índice de Desenvolvimento da Educação Básica (IDEB) consolidou a cultura de avaliação no país, pois estabeleceu de forma simples e ampla metas para o desenvolvimento das diversas unidades que compõem o sistema de ensino. Surge, portanto, uma grande oportunidade de realização de diversas pesquisas que possam identificar as melhores iniciativas na área de educação a fim de ampliar e disseminar tais práticas.

Consistente com estudos anteriores, os resultados aqui encontrados mostram que a formação dos professores de matemática teve efeitos positivos no desempenho dos estudantes do Ensino Fundamental que realizaram o Saeb em 2017, sobretudo nas regiões Sul, Sudeste e Nordeste.

Além disso, o estudo mostrou que o percentual de professores que lecionam matemática nos anos finais do ensino fundamental e que possuem formação superior adequada, avançou no período de 2014-2017. Contudo, percentual significativo de estudantes até 2017 ainda não tinham professor com essa qualificação. Nesse sentido, a política nacional de formação dos profissionais da educação, que estabelece na Meta 15 do Plano Nacional de

Educação (PNE), a universalização dos professores com formação superior, obtida em curso de licenciatura na área em que lecionam mostra-se acertada, pois, conforme evidenciado, a formação dos professores afeta positivamente o desempenho dos estudantes na etapa de ensino pesquisado. Assim, a União, os Estados, o Distrito Federal e os Municípios devem adotar todas as providências necessárias para o cumprimento da Meta 15.

Na mesma linha, o PNE estabelece que, até 2024, 50% dos professores da educação básica devem possuir formação em nível de pós-graduação. Novamente, os achados nesta pesquisa para essa variável de interesse mostram que professores com esse nível de formação afetam positivamente o desempenho dos estudantes. Razão pela qual a política nacional de formação dos professores não deve ser negligenciada. Ao contrário, os gestores devem priorizá-la dada a sua relevância para a melhoria da qualidade da educação.

7. BIBLIOGRAFIA

AARONSON, Daniel; BARROW, Lisa; SANDER, William. Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, v. 25, n. 1, p. 95-135, 2007.

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*, v.34(3), p. 527-544, 2008.

ALVES, Maria T. G.; XAVIER, Flavia, P. Indicadores multidimensionais para avaliação da infraestrutura escolar: o ensino fundamental. *Cadernos de Pesquisas*, v. 48, n. 169, 708-746, jul./set. 2018.

BARBETTA, Pedro Alberto; DE ANDRADE, Dalton Francisco; TAVARES, Heliton R. Estudo de fatores associados através de regressão quantílica hierárquica. *Estudos em Avaliação Educacional*, v. 29, n. 71, p. 320-349, 2018.

BATES, Douglas et al. Package 'lme4'. *Convergence*, v. 12, n. 1, p. 2, 2015.

BICKEL, R. (2007). *Multilevel analysis for applied research: It's just regression!* New York: The Guilford Press.

BRASIL. Lei de Diretrizes e Bases da Educação Nacional, Lei n. 9.394, de 24 de dezembro de 1996.

BRITO, M. A.; NASCIMENTO, J. S.; LIMA, J. E. Diferenciais de Desenvolvimento Econômico entre os Municípios Brasileiros: uma análise sob o prisma da regressão quantílica. *Anais do I Circuito de Debates Acadêmicos-CODE*, 2011.

CAMPOS, D. M. de S. *Psicologia da Aprendizagem*. 41ª ed. Petrópolis: Vozes, 2014

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

CHETTY, Raj; FRIEDMAN, John N.; ROCKOFF, Jonah E. Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, v. 104, n. 9, p. 2633-79, 2014.

CHIEH, Liao Yu. "Análise de desempenho acadêmico e taxa de conclusão de estudantes no decorrer da pós-graduação usando regressão quantílica e regressão logística." (2018).

- COLEMAN, J.S. et al. Equality of educational opportunity. Washington: U.S. Government Printing Office, 1966.
- COOPER, Samuel T.; COHN, Elchanan. Estimation of a frontier production function for the South Carolina educational process. *Economics of Education Review*, v. 16, n. 3, p. 313-327, 1997.
- COTTA, T. Avaliação educacional e políticas públicas: a experiência do sistema nacional de avaliação da educação básica (SAEB). *Revista do Serviço Público*. Ano 52 nº4 Brasília 2001
- DAVINO, C.; VISTOCCO, D. Quantile regression for the evaluation of student satisfaction. *Statistica Applicata*, v. 20, n. 3-4, p. 179-196, 2008.
- DE ANDRADE, Dalton Francisco; TAVARES, Heliton Ribeiro; DA CUNHA VALLE, Raquel. *Teoria da Resposta ao Item: conceitos e aplicações*. ABE, São Paulo, 2000.
- EHRENBERG, Ronald G.; BREWER, Dominic J. Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of education review*, v. 13, n. 1, p. 1-17, 1994.
- FERRÃO, M. E. (2003). *Introdução aos modelos de regressão multinível em educação*. Campinas: Editora Komedi.
- FERRÃO, Maria Eugenia et al. O SAEB – Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. *Revista Brasileira de Estudos de População*, v. 18, n. 1/2, p. 111-130, 2001.
- FIGLIO, David N., “Teacher Salaries and Teacher Quality” *Economics Letters*, 1997.
- GERACI, M. Linear quantile mixed models: The lqmm package for Laplace quantile regression. *Journal of Statistical Software*, v. 57, n. 13, p. 1-29, 2014.
- GERACI, M.; BOTTAE, M. (2014) – Linear quantile mixed models. *Statistic Computation*, v.24, p.461–479
- GERACI, Marco; BOTTAI, Matteo. Linear quantile mixed models. *Statistics and computing*, v. 24, n. 3, p. 461-479, 2014.
- GLEWWE, P.; ILIAS, N.; KREMER, M. Teacher Incentives" NBER Working Paper: No. 9671. 2003.
- GLEWWE, Paul; KREMER, Michael. Schools, teachers, and education outcomes in developing countries. *Handbook of the Economics of Education*, v. 2, p. 945-1017, 2006.

- GOLDHABER, Dan; ANTHONY, Emily. Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *The Review of Economics and Statistics*, v. 89, n. 1, p. 134-150, 2007.
- GROGGER, Jeff; EIDE, Eric. Changes in college skills and the rise in the college wage premium. *Journal of Human Resources*, p. 280-310, 1995.
- HANUSHEK, E. A. (1992). "The Trade-off Between Child Quantity and Quality." *Journal of Political Economy* 100(1): 84-117.
- HANUSHEK, Eric A. Conceptual and empirical issues in the estimation of educational production functions. *Journal of human Resources*, p. 351-388, 1979.
- HANUSHEK, Eric A. et al. Does peer ability affect student achievement?. *Journal of applied econometrics*, v. 18, n. 5, p. 527-544, 2003.
- HANUSHEK, Eric A. Interpreting recent research on schooling in developing countries. *The world bank research observer*, v. 10, n. 2, p. 227-246, 1995.
- HANUSHEK, Eric A. The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, v. 24, n. 3, p. 1141-1177, 1986.
- HANUSHEK, Eric A.; KAIN, John F.; RIVKIN, Steven G. Why public schools lose teachers. *Journal of human resources*, v. 39, n. 2, p. 326-354, 2004.
- HANUSHEK, Eric A.; RIVKIN, Steven G. Teacher quality. *Handbook of the Economics of Education*, v. 2, p. 1051-1078, 2006.
- HANUSHEK, Eric. Teacher characteristics and gains in student achievement: Estimation using micro data. *The American Economic Review*, v. 61, n. 2, p. 280-288, 1971.
- HECKMAN, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64, 605–654.
- HEYNEMAN, Stephen P., WHITE, Daphne S. (Eds.). *The quality of education and economic development*. Washington: World Bank, 1986.
- HEYNEMANN, S. The search for school effects in developing countries, 1966-86. In: Washington, DC: World Bank, Economic Development Institute, Seminar Paper. 1986.

HORTA NETO, J. Limites e possibilidades de utilização dos resultados do SAEB na gestão do sistema público de ensino: o caso do Distrito Federal. In: I Colóquio Ibero Americano de Política e Administração Escolar, 2007, Porto Alegre.

HOX, J. (2010). Multilevel analysis: Techniques and applications. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Relatório SAEB 2005. Brasília, DF INEP/MEC 2006.

JACOB, Brian A.; LEFGREN, Lars. The impact of teacher training on student achievement quasi-experimental evidence from school reform efforts in Chicago. *Journal of Human Resources*, v. 39, n. 1, p. 50-79, 2004.

KANE, Thomas J. et al. Have we identified effective teachers? Validating measures of effective teaching using random assignment. Seattle, Washington, DC: Bill & Melinda Gates Foundation. 2013.

KANE, Thomas J.; STAIGER, Douglas O. Estimating teacher impacts on student achievement: An experimental evaluation. National Bureau of Economic Research, 2008.

KLEIN, Ruben; FONTANIVE, Nilma. Avaliação em larga escala: uma proposta inovadora. *Em Aberto*, v. 15, n. 66, 1995.

KOENKER, R. (2005) – Quantile Regression. UK: Cambridge University Press, 2005.

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

KOENKER, Roger; HALLOCK, Kevin F. Quantile regression. *Journal of economic perspectives*, v. 15, n. 4, p. 143-156, 2001.

KREMER, Michael. Randomized evaluations of educational programs in developing countries: Some lessons. *American Economic Review*, v. 93, n. 2, p. 102-106, 2003.

KRUEGER, Alan, “Experimental Estimates of Education Production Functions”, *Quarterly Journal of Economics*, 1999.

KUKLA-ACEVEDO, Sharon. As características do professor são importantes? Novos resultados sobre os efeitos da preparação dos professores no desempenho dos alunos. *Revisão da Economia da Educação*, v. 28, n. 1, p. 49-57, 2009.

- LAROS, J. A., MARCIANO, J. L. P., & ANDRADE, J. M. (2010). Fatores que afetam o desempenho na prova de matemática do SAEB: Um estudo multinível. *Avaliação Psicológica*, 9(2), 173-186.
- LEE, V. L. (2008). Utilização de modelos lineares hierárquicos lineares para estudar contextos sociais: O caso dos efeitos da escola. In N. Brooke e J. F. Soares (Eds.), *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 273-296). Belo Horizonte: Editora UFMG.
- NETER, John et al. *Applied linear statistical models*. 1996.
- PINHEIRO, Jose et al. *nlme: Linear and nonlinear mixed effects models*. 2015.
- RAUDENBUSH, Stephen W.; BRYK, Anthony S. *Hierarchical linear models: Applications and data analysis methods*. sage, 2002.
- RIVKIN, Steven G.; HANUSHEK, Eric A.; KAIN, John F. Teachers, schools, and academic achievement. *Econometrica*, v. 73, n. 2, p. 417-458, 2005.
- ROCKOFF, J. The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *The American Economic Review*, 94(2), 247-25, 2004.
- SUMMERS, Anita A.; WOLFE, Barbara L. Do schools make a difference?. *The American Economic Review*, v. 67, n. 4, p. 639-652, 1977.
- WÖBMANN, Ludger. Schooling resources, educational institutions and student performance: the international evidence. *Oxford bulletin of economics and statistics*, v. 65, n. 2, p. 117-170, 2003.
- YU, Keming; ZHANG, Jin. A three-parameter asymmetric Laplace distribution and its extension. *Communications in Statistics—Theory and Methods*, v. 34, n. 9-10, p. 1867-1879, 2005.
- ZOGHBI, Ana Carolina Pereira et al. *Avaliando O Impacto Do Caderno De Apoio e Aprendizagem Na Rede De Ensino Municipal De São Paulo: Efeitos Médios e Heterogêneos*. Anais do XL Encontro Nacional de Economia. ANPEC-Associação Nacional dos Centros de Pós-graduação em Economia, 2014.

8. ANEXO

Resultados por região

Tabela 1 – Modelo região Sul

Modelo	Estadual e Municipal			Estadual			Municipal		
	Coefficiente	EP	estat-t	Coefficiente	EP	estat-t	Coefficiente	EP	estat-t
Parte Fixa									
Intercepto	136,51	1,95	70,06	141,92	2,25	63,19	110,76	3,79	29,25
Formação Adequada	0,75	0,27	2,79	0,67	0,31	2,14	0,95	0,51	1,85
Pós-Graduação	1,17	0,28	4,19	1,13	0,31	3,58	1,56	0,58	2,67
Proficiência Mat - anterior	0,52	0,00	222,96	0,52	0,00	186,84	0,54	0,00	121,50
Sexo_feminino	-7,04	0,19	-36,14	-6,97	0,23	-30,24	-7,17	0,36	-19,74
Cor_preta	-4,90	0,77	-6,37	-3,48	0,99	-3,53	-6,82	1,23	-5,55
Reprovou	-14,81	0,33	-44,61	-14,41	0,39	-36,94	-15,84	0,63	-25,08
Gosta Mat	12,00	0,21	57,16	12,09	0,25	48,86	11,84	0,40	29,86
Nse	3,84	0,15	25,82	3,97	0,17	22,79	3,49	0,28	12,25
Dep_estadual	-4,03	0,40	-10,07						
Zona_urbana	-0,38	0,72	-0,52	-0,57	0,95	-0,60	-0,51	1,09	-0,47
IN_CAPITAL	-1,43	0,73	-1,95	-0,43	0,77	-0,56	-8,91	1,99	-4,48
Infrageral	0,62	0,29	2,18	-0,64	0,33	-1,95	3,97	0,56	7,13
Equi_ap_pedag	0,35	0,12	2,89	0,52	0,15	3,47	0,20	0,21	0,97
Parte Aleatória									
Efeito Randômico - nível 1 - σ_e^2	1.019,9 (31,94)			1.024,11 (32,00)			1.008,3 (31,75)		
Variância – intercepto - nível 2 - σ_{u0}^2	107,5 (10,37)			95,65 (9,78)			123,9 (11,13)		
Correlação intraclasse	0,095			0,085			0,110		

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses.

Tabela 2 – Modelo região Sudeste

Modelo	Estadual e Municipal			Estadual			Municipal		
	Coeficiente	EP	estat-t	Coeficiente	EP	estat-t	Coeficiente	EP	estat-t
Intercepto	117,10	1,44	81,49	115,41	1,83	62,90	112,76	2,46	45,79
Formação Adequada	0,91	0,16	5,65	0,77	0,20	3,88	1,16	0,28	4,13
Pós-Graduação	3,07	0,28	10,98	4,34	0,38	11,43	1,55	0,42	3,72
Proficiência Mat - anterior	0,57	0,00	434,39	0,58	0,00	361,72	0,56	0,00	240,12
Sexo_feminino	-4,58	0,12	-38,51	-4,01	0,15	-27,43	-5,72	0,20	-28,03
Cor_preta	-3,84	0,29	-13,34	-3,93	0,37	-10,64	-3,69	0,46	-8,03
Reprovou	-13,47	0,20	-67,04	-13,04	0,26	-50,58	-14,19	0,32	-44,25
Gosta Mat	13,16	0,13	104,72	13,14	0,15	85,25	13,19	0,22	60,76
Nse	2,80	0,09	31,79	2,95	0,11	27,29	2,48	0,15	16,27
Dep_estadual	-5,38	0,27	-20,06						
Zona_urbana	-0,62	0,59	-1,05	-2,02	0,91	-2,21	-0,04	0,81	-0,05
IN_CAPITAL	-3,99	0,37	-10,86	-3,71	0,48	-7,72	-4,68	0,59	-7,93
Infrageral	1,27	0,21	6,06	0,70	0,25	2,80	2,37	0,38	6,18
Equi_ap_pedag	0,20	0,08	2,55	0,25	0,09	2,67	0,07	0,14	0,53
Parte Aleatória									
Efeito Randômico - nível 1 - σ_e^2	1.109,9 (33,32)			1.113,32 (33,36)			1.102,1 (33,20)		
Variância - intercepto - nível 2 - σ_{u0}^2	104,9 (10,24)			94,96 (9,74)			124,8 (11,17)		
Correlação intraclasse	0,086			0,078			0,101		

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses.

Tabela 3 – Modelo região Centro-Oeste

Modelo	Estadual e Municipal			Estadual			Municipal		
	Coefficiente	EP	estat-t	Coefficiente	EP	estat-t	Coefficiente	EP	estat-t
Parte Fixa									
Intercepto	119,91	2,83	42,32	126,36	3,41	37,09	106,53	5,05	21,10
Formação Adequada	0,54	0,38	1,42	0,11	0,46	0,24	1,66	0,70	2,38
Pós-Graduação	1,65	0,58	2,84	1,19	0,73	1,65	2,61	0,96	2,73
Proficiência Mat - anterior	0,56	0,00	165,71	0,55	0,00	140,09	0,57	0,01	88,99
Sexo_feminino	-6,78	0,29	-23,75	-6,58	0,33	-19,66	-7,32	0,54	-13,43
Cor_preta	-2,71	1,02	-2,65	-3,10	1,20	-2,58	-1,47	1,96	-0,75
Reprovou	-13,29	0,42	-31,42	-13,96	0,49	-28,23	-11,27	0,81	-13,86
Gosta Mat	12,92	0,30	42,43	13,37	0,36	37,60	11,52	0,59	19,59
Nse	2,93	0,21	14,16	2,76	0,24	11,49	3,47	0,40	8,57
Dep_estadual	1,14	0,59	1,94						
Zona_urbana	1,62	0,95	1,71	2,53	1,37	1,84	-0,88	1,32	-0,67
IN_CAPITAL	-5,96	0,66	-9,02	-6,97	0,77	-9,08	-3,34	1,28	-2,62
Infrageral	1,66	0,44	3,77	1,03	0,52	1,99	3,39	0,82	4,14
Equi_ap_pedag	0,12	0,17	0,74	-0,07	0,19	-0,36	0,39	0,31	1,26
Parte Aleatória									
Efeito Randômico - nível 1 - σ_e^2	1.072,87 (32,75)			1.098,13 (33,14)			998,58 (31,60)		
Variância – intercepto - nível 2 - σ_{u0}^2	85,65 (9,25)			83,73 (9,15)			83,57 (9,14)		
Correlação intraclasses	0,073			0,070			0,077		

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses.

Tabela 4 – Modelo região Nordeste

Modelo	Estadual e Municipal			Estadual			Municipal		
	Coeficiente	EP	estat-t	Coeficiente	EP	estat-t	Coeficiente	EP	estat-t
Parte Fixa									
Intercepto	121,92	1,74	70,06	126,34	3,58	35,26	120,16	2,00	60,07
Formação Adequada	0,37	0,29	1,27	1,00	0,57	1,75	0,10	0,34	0,28
Pós-Graduação	0,40	0,34	1,17	0,41	0,70	0,60	0,39	0,39	1,01
Proficiência Mat - anterior	0,53	0,00	272,14	0,53	0,00	116,40	0,53	0,00	245,68
Sexo_feminino	-9,03	0,17	-53,57	-8,61	0,39	-22,33	-9,12	0,19	-48,72
Cor_preta	-2,17	0,50	-4,37	-2,76	0,98	-2,81	-1,99	0,57	-3,47
Reprovou	-11,53	0,23	-49,75	-10,42	0,52	-19,90	-11,78	0,26	-45,63
Gosta Mat	12,13	0,19	65,55	11,39	0,40	28,27	12,34	0,21	59,26
Nse	2,04	0,12	16,97	2,00	0,27	7,28	2,05	0,13	15,38
Dep_estadual	-4,07	0,46	-8,77						
Zona_urbana	-2,16	0,42	-5,16	-1,31	1,35	-0,97	-2,36	0,45	-5,24
IN_CAPITAL	-2,88	0,61	-4,70	-4,94	0,71	-6,97	-0,59	0,95	-0,62
Infrageral	2,70	0,30	8,96	2,06	0,56	3,70	2,81	0,35	8,05
Equi_ap_pedag	0,34	0,10	3,28	-0,28	0,18	-1,59	0,51	0,12	4,13
Parte Aleatória									
Efeito Randômico - nível 1 - σ_e^2	1.104,4 (33,23)			1.087,4 (32,98)			1.108,7 (33,30)		
Variância – intercepto - nível 2 - σ_{u0}^2	183,8 (13,56)			108,5 (10,42)			201,2 (14,19)		
Correlação intraclasse	0,142			0,090			0,153		

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses.

Tabela 5 – Modelo região Norte

Modelo	Estadual e Municipal			Estadual			Municipal		
	Coefficiente	EP	estat-t	Coefficiente	EP	estat-t	Coefficiente	EP	estat-t
Parte Fixa									
Intercepto	119,40	2,07	57,74	114,41	3,14	36,38	125,78	2,90	43,36
Formação Adequada	0,05	0,32	0,14	0,04	0,42	0,10	-0,06	0,49	-0,12
Pós-Graduação	0,59	0,61	0,96	-0,52	0,77	-0,68	2,61	1,00	2,60
Proficiência Mat - anterior	0,50	0,00	175,81	0,52	0,00	138,69	0,48	0,00	108,25
Sexo_feminino	-8,78	0,24	-36,50	-8,49	0,32	-26,78	-9,17	0,37	-24,90
Cor_preta	-0,94	0,92	-1,02	-2,38	1,23	-1,94	0,83	1,39	0,60
Reprovou	-10,94	0,32	-34,12	-11,29	0,45	-25,13	-10,73	0,46	-23,43
Gosta Mat	11,75	0,27	43,58	11,80	0,35	33,43	11,53	0,42	27,65
Nse	2,16	0,17	12,71	2,33	0,22	10,49	1,92	0,26	7,28
Dep_estadual	-2,17	0,52	-4,15						
Zona_urbana	-1,14	0,64	-1,78	0,04	1,24	0,03	-1,61	0,76	-2,12
IN_CAPITAL	0,33	0,60	0,56	-2,67	0,72	-3,69	7,75	1,07	7,24
Infrageral	3,58	0,37	9,59	3,52	0,53	6,71	3,34	0,52	6,38
Equi_ap_pedag	0,45	0,14	3,21	0,40	0,20	2,02	0,26	0,20	1,30
Parte Aleatória									
Efeito Randômico - nível 1 - σ_e^2	1.064 (32,62)			1.057,1 (32,51)			1.071,65 (32,73)		
Variância – intercepto - nível 2 - σ_{u0}^2	103,7 (10,18)			107,1 (10,35)			92,61 (9,60)		
Correlação intraclasse	0,088			0,092			0,080		

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses.

Tabela 6 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, na região Sul.

Escolas das redes estadual e municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,70 (0,64)	0,98 ** (0,36)	1,01 ** (0,30)	1,14 ** (0,37)	1,27 * (0,48)
Pós-Graduação	1,23 . (0,65)	0,81 * (0,39)	0,53 (0,37)	0,55 (0,43)	1,71 * (0,66)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede estadual					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	1,30 . (0,65)	0,45 (0,37)	0,82 * (0,41)	0,86 * (0,41)	0,79 (0,58)
Pós-Graduação	0,14 (0,80)	1,07 * (0,46)	0,81 * (0,38)	0,91 * (0,36)	0,99 (0,64)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,59 (1,02)	-0,08 (0,64)	0,90 (0,59)	0,77 (0,63)	1,55 (0,97)
Pós-Graduação	2,56 * (1,21)	1,27 (0,84)	1,17 . (0,66)	1,14 (0,80)	0,36 (1,12)
Controles: Alunos e escolas, conforme quadro 1.					

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses. Significância: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’ 1.

Tabela 7 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, na região Sudeste.

Escolas das redes estadual e municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,99 *	0,94 **	0,82 ***	1,16 ***	0,69 *
	(0,41)	(0,29)	(0,23)	(0,24)	(0,39)
Pós-Graduação	3,72 ***	3,72 ***	3,74 ***	3,65 ***	3,33 ***
	(0,84)	(0,51)	(0,30)	(0,49)	(0,61)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede estadual					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,78 .	0,72 .	0,74 **	0,72 **	0,40
	(0,43)	(0,36)	(0,22)	(0,27)	(0,47)
Pós-Graduação	5,47 ***	4,24 ***	4,24 ***	3,83 ***	3,73 ***
	(1,08)	(0,63)	(0,58)	(0,65)	(0,90)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	1,03	0,95 .	1,23 **	1,11 *	0,56
	(0,63)	(0,49)	(0,39)	(0,42)	(0,58)
Pós-Graduação	0,79	1,29 *	1,59 **	2,02 **	1,72 .
	(0,88)	(0,63)	(0,51)	(0,63)	(0,94)
Controles: Alunos e escolas, conforme quadro 1.					

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses. Significância: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’ 1.

Tabela 8 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, na região Centro-Oeste.

Escolas das redes estadual e municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,31 (0,75)	0,05 (0,44)	0,45 (0,39)	0,39 (0,53)	0,91 (0,68)
Pós-Graduação	0,72 (1,23)	2,36 * (0,93)	1,57 * (0,70)	1,88 ** (0,65)	2,17 . (1,14)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede estadual					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,23 (0,77)	0,35 (0,64)	0,27 (0,54)	0,17 (0,65)	0,17 (0,91)
Pós-Graduação	1,57 (1,60)	2,18 * (0,92)	1,01 (0,83)	1,24 (0,81)	2,33 * (1,15)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	1,01 (1,45)	0,93 (0,84)	1,54 * (0,75)	1,48 . (0,76)	2,27 . (1,31)
Pós-Graduação	2,25 (1,81)	2,79 . (1,58)	2,74 * (1,10)	2,64 * (1,08)	1,80 (1,46)
Controles: Alunos e escolas, conforme quadro 1.					

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses. Significância: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’ 1.

Tabela 9 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, na região Nordeste.

Escolas das redes estadual e municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,91 (0,69)	0,72 . (0,42)	-0,09 (0,37)	-0,05 (0,40)	-0,05 (0,55)
Pós-Graduação	1,48* (0,61)	0,87 (0,52)	0,64 (0,44)	0,67 (0,51)	0,07 (0,62)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede estadual					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	2,01 (1,39)	1,24 (0,75)	0,86 (0,65)	1,06 (0,71)	-0,07 (1,06)
Pós-Graduação	0,33 (1,79)	2,27 * (1,07)	0,58 (0,80)	0,37 (0,95)	1,96 (1,31)
Controles: Alunos e escolas, conforme quadro 1.					
Apenas escolas da rede municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,98 (0,69)	0,27 * (0,53)	-0,46 (0,37)	-0,15 (0,44)	-0,26 (0,57)
Pós-Graduação	0,70 (0,78)	0,82 (0,68)	0,79 (0,52)	0,25 (0,53)	-0,26 (0,92)
Controles: Alunos e escolas, conforme quadro 1.					

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses. Significância: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1.

Tabela 10 – Efeitos da formação do professor sobre o desempenho dos estudantes do 9º ano em matemática, na região Norte.

Escolas das redes estadual e municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	0,46 (0,80)	0,22 (0,59)	-0,40 (0,43)	-0,36 (0,43)	0,33 (0,61)
Pós-Graduação	0,77 (1,40)	-0,04 (0,87)	0,38 (0,82)	1,14 (0,67)	1,68 (1,13)
Controles: Alunos e Escolas					
Apenas escolas da rede estadual					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	-0,12 (1,01)	0,15 (0,61)	-0,22 (0,54)	-0,92 (0,70)	1,16 (0,87)
Pós-Graduação	-0,41 (1,61)	-1,79 (1,01)	-0,33 (0,81)	1,09 (0,89)	0,26 (1,19)
Controles: Alunos e Escolas					
Apenas escolas da rede municipal					
Variáveis	Quantil				
	10%	25%	50%	75%	90%
Formação Adequada	2,26 * (1,24)	-0,80 (0,70)	-0,64 (0,50)	-0,08 (0,59)	1,00 (0,82)
Pós-Graduação	3,69 (2,53)	1,92 (1,78)	1,80 (1,45)	2,85 * (1,30)	4,97 ** (1,67)
Controles: Alunos e Escolas					

Fonte: Elaboração própria.

Nota: Erro padrão entre parênteses. Significância: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1.