

O EFEITO-ESCOLA E O DESEMPENHO DOS ESTUDANTES DO 9º ANO DO ENSINO FUNDAMENTAL EM ESCOLAS PÚBLICAS LOCALIZADAS NO CAMPO¹

Robson dos Santos²

RESUMO

O trabalho analisou o desempenho em Língua Portuguesa e em Matemática dos estudantes do 9º ano do ensino fundamental em estabelecimentos públicos (municipais e estaduais) localizados nas áreas rurais do Brasil. Para isso utilizou uma amostra de dados do Sistema de Avaliação da Educação Básica (SAEB) aplicado em 2017, e investigou quais características das escolas no campo e dos estudantes estão correlacionadas à proficiência obtida. Para isso, foram utilizadas estatísticas descritivas, modelos de regressão linear e modelos multinível. Os resultados apontam que os estudantes em escolas no campo que trabalham fora e os que tinham sido reprovados apresentavam desempenho inferior em Língua Portuguesa e Matemática. A escolaridade dos pais e o nível socioeconômico apresentavam associação positiva com o desempenho, assim como a taxa de participação dos estudantes da escola na avaliação. Além disso, os dados sugerem que a escola no campo apresenta efeito sobre o aprendizado, o que precisa ser analisado com mais detalhes e com dados adicionais.

Palavras-chave: Efeito-Escola. Fatores associados. Avaliação. Escolas no campo. Desigualdades.

1 INTRODUÇÃO

A compreensão dos resultados obtidos pelos estudantes em avaliações externas de larga escala, como as realizadas pelo SAEB, precisa considerar, além de aspectos cognitivos e das habilidades individuais, os vários contextos e condições em que o aprendizado e o acúmulo de conhecimentos ocorre. Assim, a análise dos fatores associados ao desempenho escolar, seja os fatores relacionados à escola ou os associados ao estudante e sua família, constitui uma área de investigação que conta com ampla literatura especializada.

Dando sequência a uma tradição de debates que se inicia, sobretudo, a partir das conclusões expostas pelo Relatório Coleman (BROOKE; SOARES, 2008), que apontava um efeito reduzido da escola frente às condições socioeconômicas dos estudantes (KARINO; LAROS, 2017), as pesquisas têm buscado delimitar o papel específico da instituição escolar no desempenho cognitivo dos estudantes. Para isso, considera-se a estrutura da escola, o currículo e a formação de professores, mas também fatores como a origem social do

¹ Artigo apresentado como requisito parcial para a conclusão de cursos de Especialização em Estatísticas e Avaliação Educacional.

² Sociólogo. Pesquisador-Tecnologista em Informações e Avaliações Educacionais (INEP).

estudante, seu nível socioeconômico e o capital cultural familiar como elementos importantes para se explicar a proficiência dos indivíduos em testes de larga escala. Isso porque as habilidades que precisam ser mobilizadas para o aprendizado não se constituem apenas no espaço de socialização escolar, mas derivam também de oportunidades e práticas que ocorrem na comunidade, na família, em espaços culturais não escolares que se distribuem de forma desigual.

A área de estudos sobre efeito-escola, posicionada em um contexto de investigação sobre a qualidade educacional (BROOKE; SOARES, 2008), engloba análises que se dedicam a entender os fatores associados à proficiência; compreender qual parte da variação no desempenho alcançado pelo estudante pode ser explicada especificamente a partir do trabalho feito na/pela escola e qual deriva da origem e das condições socioeconômicas dos estudantes, isto é, que são externas à escola e ao seu trabalho, mas que exercem uma função socializadora relevante para a incorporação de habilidades e conhecimentos, de um capital cultural valorizado pelo currículo escolar. Junto a isso, as investigações buscam também aprender se a escola é o mais eficaz possível em garantir o aprendizado de forma equitativa independente da origem e das condições socioeconômicas do estudante.

O tema comporta diversas complexidades, afinal interno e externo são dimensões que se intercambiam no trabalho escolar, dado que os processos de socialização, aprendizado e aquisição de conhecimentos não se circunscrevem a um único espaço social ou institucional, o que exige refinamento teórico e metodológico das análises. Disso advém a importância do levantamento de dados contextuais sobre a escola e os estudantes para compreender e interpretar os resultados. Isso inclusive ganha relevância no campo da avaliação educacional e tem fomentado diversos estudos sobre o contexto socioeconômico da escola e dos estudantes.

De todo modo, é fundamental para as políticas educacionais compreender se e como aquilo que é feito pela escola, a formação dos professores, aspectos da estrutura escolar são capazes de exercer um efeito positivo sobre o aprendizado dos estudantes. Mas, dada a relevância dos fatores contextuais, é fato que isso pode variar muito em função da etapa, da série ou do tipo de conhecimento focado (ANDRADE; SOARES, 2008). Nessa direção, o local ou território em que a escola se posiciona também representa um componente importante para refletir sobre seus possíveis efeitos sobre o aprendizado. No Brasil, em específico, as áreas rurais constituem um espaço diferenciado de socialização – em termos de oportunidades, relações, equipamentos culturais etc. – e que precisam ser compreendidas com mais detalhes nos estudos sobre o efeito-escola e/ou os fatores contextuais relacionados à proficiência nas avaliações em larga escala.

No caso brasileiro, a principal fonte de informações utilizada para realizar as análises são os dados provenientes do Sistema Nacional de Avaliação da Educação Básica (SAEB), sob responsabilidade do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Por meio dos instrumentos do SAEB, que congregam provas para aferir a proficiência em Língua Portuguesa e Matemática e questionários socioeconômicos, os estudantes do 5º e 9º ano do ensino fundamental e os da 3ª série do ensino médio de escolas públicas que tem pelo menos 10 alunos na referida série são avaliados a cada dois anos; também são aplicados questionários aos professores e diretores de escola. Como o SAEB emprega a Teoria de Resposta ao Item (TRI), as edições podem ter seus resultados comparados, possibilitando, assim, que se disponha de um quadro acerca dos avanços e/ou dos recuos nos sistemas de ensino.

Com os dados do SAEB, inúmeras investigações têm se voltado para compreender o efeito-escola e os fatores contextuais, considerando diversas características e desagregações, inclusive por estados e municípios (SOARES; ALVES, 2013). Isso tem fornecido evidências para um melhor entendimento acerca da qualidade da educação no país e, sobretudo, sobre as desigualdades que a caracterizam em termos de localização, estrutura, dependência, público atendido, entre outras dimensões relevantes.

Nesse contexto, o objeto desse estudo é analisar o desempenho em Língua Portuguesa e em Matemática dos estudantes do 9º ano do ensino fundamental em escolas públicas (estaduais e municipais) localizadas em áreas rurais, com base nos dados do SAEB (2017). É fato que a qualidade educacional e os resultados alcançados pelo trabalho escolar não se restringem à proficiência obtida na avaliação. Entretanto, eles constituem medidas importantes sobre a qualidade da educação oferecida, daí a relevância de recorrer a elas.

É importante registrar que, atualmente, rural e urbano não são mais entendidos como polos antagônicos marcados por diferenças extremas, mas como contínuos de sociabilidade em permanente interação (FAVARETO, 2014). As formas de produção e circulação no campo passaram por relevantes transformações (VEIGA, 2006). Apesar das mudanças ocorridas, ainda existem diferenças relevantes de estrutura, assim como de acesso, rendimento e desempenho médio entre as escolas nas áreas rurais e nas áreas urbanas. Por esse motivo, é importante analisar se o que a literatura tem apontado como efeito-escola e os fatores contextuais preserva a mesma magnitude quando se considera em específico as escolas localizadas no campo.

A investigação também se justifica na medida em que, nas últimas décadas, políticas educacionais específicas para as populações do campo foram implantadas visando melhorar a

qualidade da oferta e o aprendizado, como é o caso do Programa Nacional de Educação do Campo (PRONACAMPO), o Programa Nacional de Educação na Reforma Agrária (PRONERA), o Programa Nacional do Livro Didático – Campo (PNLD – Campo), o Programa de Apoio à Formação Superior em Licenciatura em Educação do Campo (PROCAMPO), o Plano Nacional de Educação (PNE), entre outras ações (SANTOS, 2019).

Essas políticas ocorrem em associação à emergência do paradigma da educação do campo, isto é, de uma proposta de organização pedagógica da escola no qual os espaços de aprendizado e os espaços de reprodução das condições de vida, da cultura e das relações sociais não se distinguem. Esse paradigma parte “dos diferentes sujeitos do campo, do seu contexto, sua cultura e seus valores, sua maneira de ver e de se relacionar com o tempo, a terra, com o meio ambiente” (ARROYO; CALDART; MOLINA, 2011, p.14) Assim, é importante compreender com mais detalhes o papel que a escola localizada nas áreas rurais ocupa no aprendizado de seus estudantes considerando as suas particularidades socioeconômicas e as condições em que o direito à educação se concretiza no campo.

A escolha do 9º ano se deu em função de ser quando se encerra a etapa de ensino fundamental, que comporta, para os estudantes em análise, longo período de interação e socialização em escolas localizadas no campo. Além disso, ainda é frequente que, após a conclusão do ensino fundamental, a maior parte dos estudantes seja deslocada para realizar o ensino médio em estabelecimentos localizados nas áreas urbanas, o que, especialmente no caso dos estudantes que residem no campo, é acompanhado de diversas modificações socioculturais e que podem ter efeitos sobre a aprendizagem. Todavia, investigar tais transições foge ao escopo desse trabalho.

Assim, considerando as condições sociais específicas das áreas rurais brasileiras, com seu acesso diferenciado a serviços e bens culturais, uma estrutura ocupacional particular, particularidades na composição demográfica, uma escolaridade média menor, diferenças nas formas de inserção comunitária da escola, entre outras questões (SANTOS, 2019), o objetivo do trabalho foi compreender as variações no desempenho dos estudantes em Língua Portuguesa e Matemática em função das características da escola: formação dos professores, existência de bibliotecas, internet e o nível socioeconômico médio; e dos estudantes: posse de bens, escolaridade dos pais, práticas culturais e de estudo, trabalho fora de casa e reprovação prévia. Em termos gerais, as perguntas para as quais o trabalho buscou algumas respostas são: como os fatores internos à escola e as condições dos estudantes se correlacionam ao desempenho alcançado; e se o efeito-escola no campo assumiria magnitudes distintas quando

cotejado com aqueles encontrados por outros estudos dedicados a análise das escolas como um todo.

O texto possui a seguinte subdivisão: além dessa introdução, na primeira parte do desenvolvimento é feita uma breve revisão dos estudos sobre efeito-escola e sobre fatores contextuais buscando destacar os métodos empregados e os principais achados, com ênfase nos trabalhos que utilizam o SAEB; na sequência é feita uma apresentação descritiva das escolas públicas de ensino fundamental com oferta de 9º ano nas áreas rurais e das características de seus estudantes, a partir dos dados do questionário do SAEB de 2017; na segunda parte do desenvolvimento são feitas análises descritivas das variáveis empregadas nos modelos; depois são apresentados os modelos e a análise dos resultados; por fim, são destacadas as principais conclusões.

2 DESENVOLVIMENTO

A análise sobre a eficácia do trabalho da escola nas áreas rurais deve considerar a existência de uma ampla literatura sobre efeito-escola e fatores associados, com estudos provenientes da educação, economia, psicologia, sociologia, entre outras áreas do conhecimento. Não é possível realizar uma revisão detalhada da literatura na área, motivo pelo qual se apresentam os achados e reflexões de alguns artigos relevantes de modo que se tenha uma ideia geral das contribuições trazidas e se sistematizem algumas referências teóricas e metodológicas para o estudo. De todo modo, essa breve revisão é relevante para uma compreensão e análise dos dados e dos modelos posteriormente desenvolvidos.

2.1 O efeito-escola e os fatores contextuais

A ampliação do acesso à educação básica em países da Europa e nos Estados Unidos foi seguida, na segunda metade do século XX, por uma preocupação crescente em compreender se a escola era efetiva enquanto instituição dedicada à aprendizagem e se essa efetividade era igualmente distribuída na população. A partir da década de 1960, as análises buscaram investigar de forma empírica se a instituição escolar cumpria o papel democratizante e republicano que sua expansão almejava, ou se reproduzia e legitimava desigualdades sociais previamente existentes (BOURDIEU, 2003). Assim, o desenvolvimento de análises sobre o efeito-escola se dá em paralelo às pesquisas que buscam ressaltar os aspectos sociais, econômicos e culturais que contextualizam o processo educacional e que tem relação, entre outras dimensões, com variações no desempenho.

Entre os vários estudos que desde então apareceram, o chamado Relatório Coleman, coordenado pelo sociólogo James Coleman, trouxe questionamentos acerca do efeito equitativo da escola e subsidiou importantes debates no campo das políticas educacionais (BROOKE; SOARES, 2008). Entre os achados do referido Relatório, ganhou relevância os que indicavam efeito reduzido da escola sobre o aprendizado quando os controles por nível socioeconômico eram introduzidos nos modelos utilizados. Como explica o autor:

A primeira descoberta é que as escolas são extraordinariamente similares no efeito que elas exercem no desempenho de seus alunos, quando o background socioeconômico é levado em consideração. É sabido que fatores socioeconômicos possuem uma grande relação com o desempenho acadêmico. Entretanto, quando esses fatores são controlados estatisticamente, parece que as diferenças entre as escolas são responsáveis somente por uma pequena fração das diferenças no desempenho dos alunos (COLEMAN, 2008, p. 29).

Esse trabalho era focado no sistema de ensino americano e na análise de seus potenciais para promover a equidade em uma sociedade clivada por desigualdades. Desde então, um conjunto de estudos se acumulou em diversos países, buscando compreender as relações do sistema de ensino com as desigualdades, assim como o efeito da escola no aprendizado, considerando as características dos estudantes. No caso brasileiro, esses estudos passam a ganhar mais projeção a partir da década de 1990, com a consolidação do SAEB (ALVES; SOARES, 2007).

Em revisão da literatura sobre a temática, Karino e Laros (2017) destacam a relevância do trabalho coordenado por Coleman e elencam vários estudos posteriores que se desenvolveram sobre a eficácia da escola em âmbito internacional. No caso brasileiro, os autores também destacam o desenvolvimento da temática, sobretudo a partir da consolidação do SAEB. Na literatura nacional, assim como na internacional, os estudos “têm apontado para um conjunto de fatores (contextuais, organizacionais, de monitoramento e pedagógicos) que tornam a escola eficaz ao contribuir com o desenvolvimento do estudante” (KARINO; LAROS, 2017, p. 117). Ou seja, as pesquisas têm encontrado evidências de que características específicas da escola podem reduzir as desigualdades de aprendizagem. Entretanto, alertam os autores a partir da revisão:

Os resultados estrangeiros e brasileiros têm consistentemente apontado que há desigualdade entre estudantes por nível socioeconômico, raça e gênero. Mais desanimador ainda são os resultados dos estudos sobre equidade que mostram que, ao se promover eficácia escolar, se promove um aumento das desigualdades, uma vez que estudantes mais favorecidos tendem a usufruir

mais das melhorias no ambiente escolar. Debruçar-se sobre a problemática de como alcançar um sistema educacional mais igual e equânime é um desafio urgente (KARINO; LAROS, 2017, p. 119).

No caso brasileiro, as desigualdades assumem recortes socioeconômicos, territoriais e regionais que adicionam complexidades à reflexão sobre o papel dos estabelecimentos de ensino na garantia de um aprendizado mais equitativo. Uma questão central que se coloca é se a escola e o trabalho nela desenvolvido são eficazes em garantir o aprendizado dos estudantes em todos os contextos. Como destacam Andrade e Soares (2008), os possíveis efeitos da escola são múltiplos, mas, de modo usual, “o conceito de ‘efeito da escola’ é utilizado para medir a capacidade de as escolas, por meio de seu projeto pedagógico e de suas políticas internas, influenciarem o desempenho cognitivo de seus alunos” (p. 381). Todavia, é preciso compreender se tal capacidade se mantém em todos os espaços, localizações e com os mais diversos públicos.

Como dito antes, os resultados educacionais, incluso os de aprendizagem, decorrem de diversos fatores que estão relacionados tanto ao contexto interno da escola quanto às dimensões sociais, culturais e econômicas que caracterizam os estudantes e o ambiente em que a escola se insere. É fato, como explicam Ferrão e Fernandes (2003), que os fatores escolares suscitam maior preocupação, pois se referem a dimensões de possível intervenção por parte das políticas públicas, especialmente daquelas sob a gestão da escola ou dos sistemas de ensino. Porém, advertem os autores:

Não devemos, entretanto, esquecer-nos de que a contribuição da escola no desenvolvimento intelectual e social das crianças está condicionada às características extra-escolares que estas carregam: as características sociais, econômicas e culturais de sua família e de seu ambiente cotidiano, e o conhecimento que os alunos já têm quando entram na escola (FERRÃO; FERNANDES, 2003, p. 01).

As análises sobre o desempenho dos estudantes nas avaliações em larga escala a partir dos fatores contextuais associados, do efeito-escola e/ou da eficácia escolar contam com inúmeras investigações no Brasil, das quais uma revisão sistemática escapa ao escopo desse trabalho. Visando produzir uma síntese, um aspecto comum às várias análises é o esforço de tentar compreender as relações que o processo de escolarização possui com as diversas desigualdades (PICANÇO; MORAIS, 2016).

Nesse sentido, a diferença de desempenho por localização de residência, que é o foco desse trabalho, é um fenômeno que precisa ser compreendido e analisado. De um lado, pode estar associado ao nível socioeconômico menor dos residentes no campo no Brasil, que

implica, por exemplo, menor acesso a bens culturais e menor escolaridade dos pais, apontados pela literatura como aspectos associados ao desempenho. Por outro lado, a questão da qualidade educacional das escolas no campo precisa ser considerada: questões como a ausência de instituições próximas, a limitação na infraestrutura (SOARES NETO; JESUS; KARINO; ANDRADE, 2013), nos recursos didático-pedagógicos, bem como a formação dos professores podem ser elementos importantes para compreender a situação, dada a necessidade de “redução das intensas desigualdades e da precariedade do acesso à educação escolar, nos diferentes níveis, no meio rural” (MOLINA; FREITAS, 2011, p. 19). Em paralelo, a ausência de um projeto pedagógico que reflita as configurações e a realidade do campo também pode cumprir papel relevante.

Para compreender melhor tais questões, é analisado na sequência o desempenho dos estudantes do 9º ano do ensino fundamental em escolas públicas localizadas no campo, com base nos dados do SAEB de 2017. O que se busca é compreender quais são os fatores associados ao desempenho de estudantes nos estabelecimentos localizados nas áreas rurais e como o efeito-escola pode ser observado nesse contexto.

2.2 Dados utilizados

A base de dados utilizada nesse trabalho, como dito, foi a disponibilizada pelo Sistema de Avaliação da Educação Básica (SAEB), relativo ao ano de 2017³. O INEP, responsável pela referida avaliação, publica o arquivo de microdados contendo resultados da avaliação para o 5º e o 9º ano do ensino fundamental e para o 3º ano do ensino médio, com foco em Língua Portuguesa e em Matemática. Além dos resultados nos testes também estão disponíveis os dados do questionário socioeconômico dos estudantes, sobre a escola, os docentes e diretores e a documentação necessária para compreender a avaliação.

O trabalho utilizou os resultados da avaliação em Língua Portuguesa e Matemática, os dados do questionário socioeconômico com as características dos estudantes do 9º ano do ensino fundamental em escolas públicas localizadas no campo, bem como algumas variáveis sobre o estabelecimento de ensino, como o percentual de professores com formação adequada, condições de infraestrutura e o nível socioeconômico médio da escola.

Os microdados relativos aos estudantes do 9º ano foram juntados aos dados com informações sobre as escolas. Na base de estudantes, foram mantidos apenas os casos em que o estudante possuía proficiência calculada em Língua Portuguesa e Matemática e que havia

³ Os microdados foram publicados no ano de 2018 (INEP, 2018).

preenchido o questionário socioeconômico. A base de escolas, por sua vez, manteve apenas as escolas estaduais e municipais localizadas no campo, com matrículas no 9º ano do ensino fundamental; que tinham dados sobre a nota média em Língua Portuguesa (LP) e Matemática (MT); que apresentavam consistência com o Censo Escolar; que tinham valores calculados para a taxa de participação na prova e que dispunham de informações para o indicador de adequação da formação docente. É importante destacar que o Indicador de Nível Socioeconômico da Escola, que se trata de uma média dos escores calculados com dados para os estudantes, não dispunha de informações relativas ao ano de 2017, motivo pelo qual utilizou-se o indicador referente ao ano de 2015.

Após os tratamentos e a junção das bases, a amostra final do SAEB 2017 que foi utilizada no trabalho era composta por 180.408 estudantes, distribuídos por 8.415 escolas públicas (estaduais e municipais) localizadas nas áreas rurais do Brasil.

2.3 Análises descritivas

Antes de analisar com mais detalhes os dados referentes especificamente aos estabelecimentos nas áreas rurais, foco desse trabalho, é relevante apresentar algumas informações gerais do SAEB de 2017 para o 9º ano do ensino fundamental, considerando todas as escolas estaduais e municipais (urbanas e rurais).

De acordo com as tabelas de resultados gerais divulgadas pelo INEP junto aos microdados do SAEB de 2017, em LP⁴, a média entre as escolas no campo era de 235,4 e entre as urbanas de 253,7, o que as posicionava em faixas distintas da escala de proficiência (INEP, 2018)⁵. Situação similar era observada no caso de MT⁶, na qual a proficiência média era superior entre as escolas localizadas nas áreas urbanas, 251,8, frente às rurais, 235,3 (INEP, 2018).

Quando se observa a situação por Estados nota-se que, em quase todos, a média em Língua Portuguesa e em Matemática das escolas rurais era inferior ao observado entre as escolas localizadas nas áreas urbanas (Gráfico 1).

Gráfico 1 – Proficiência média em Língua Portuguesa e em Matemática nas escolas públicas (estaduais e municipais) no 9º ano do ensino fundamental, por localização e UF (2017)

⁴ Proficiência do aluno em língua Portuguesa transformada na escala única do SAEB, com média = 250, desvio = 50 (do SAEB/97) (SAEB, 2017).

⁵ O material que acompanha a documentação do SAEB contém as escalas.

⁶ Proficiência do aluno em Matemática transformada na escala única do SAEB, com média = 250, desvio = 50 (do SAEB/97). (SAEB, 2017).

verifica a décadas e não se explicita apenas no desempenho em avaliações em larga escala, mas se refere também ao acesso e estrutura, entre outras dimensões (BOF et al, 2007; SOARES NETO et al, 2013; MOLINA; FREITAS, 2011).

Feita essa breve comparação a partir das médias das escolas públicas em nível estadual, cabe agora uma análise mais detalhada sobre as variáveis e o conjunto de estudantes e escolas no campo inclusos na amostra aqui analisada. Como dito antes, o conjunto de dados utilizados, após os tratamentos e as opções de exclusão descritas, engloba 180.408 estudantes distribuídos por 8.415 escolas localizadas no campo em todos os estados do Brasil. É sobre esse conjunto de dados que as análises a seguir se debruçam.

É importante verificar, inicialmente, como as proficiências médias em Língua Portuguesa e em Matemática dos estudantes variam em função das características das escolas em que estudam. Como se depreende dos dados da amostra, naqueles estabelecimentos localizados no campo em que existia biblioteca, laboratório de ciências e acesso à internet, as médias dos estudantes eram, no geral, maiores quando comparadas às médias dos estudantes em estabelecimentos que não possuíam tais equipamentos (Tabela 01). Essa informação descritiva sugere a importância da infraestrutura das escolas como um componente para se compreender o desempenho dos estudantes.

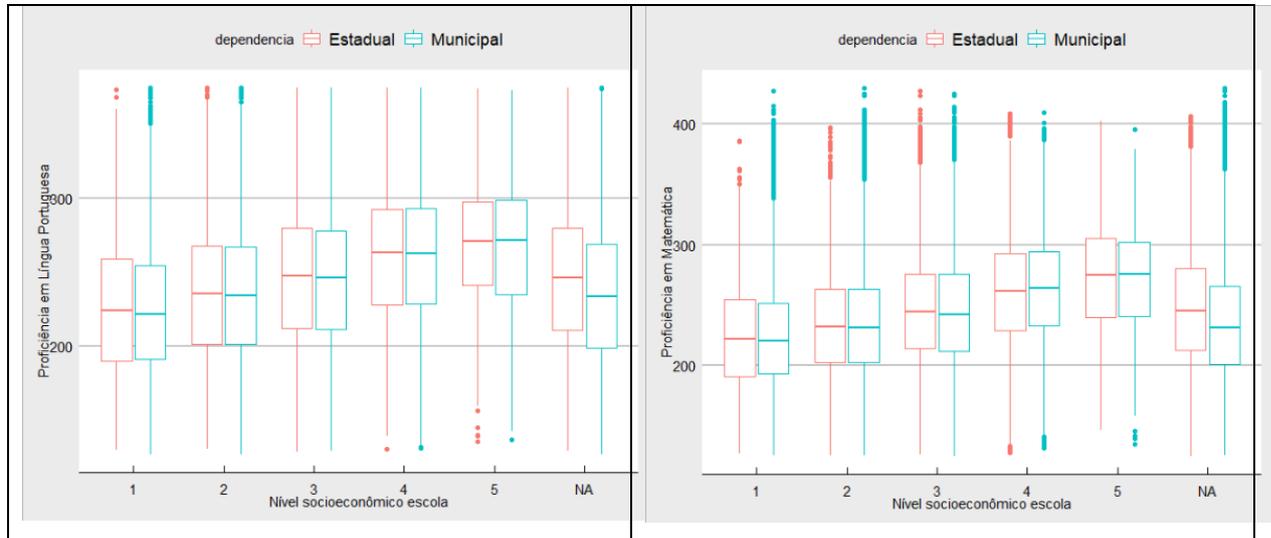
Tabela 1 – Média de proficiência dos estudantes em Língua Portuguesa e Matemática para o 9º ano do ensino fundamental em escolas localizada no campo, por existência de equipamentos – Brasil (2017).

Equipamento	Situação	Média em Língua Portuguesa 9º ano	Média em Matemática 9º ano
Biblioteca	Possui	242,0	242,4
	Não Possui	230,2	229,5
Laboratório Ciências	Possui	246,9	247,8
	Não Possui	236,8	236,6
Internet	Possui	242,3	242,5
	Não Possui	232,9	232,7

Fonte: INEP/SAEB (2017). Elaboração própria.

Para compreender a variação da proficiência é importante realizar também a comparação a partir do nível socioeconômico e da dependência administrativa, que serão inclusas nos modelos como variáveis de controle. Nesse caso, existe evidência de uma correlação entre as proficiências médias em Língua Portuguesa e em Matemática, que aumenta conforme se amplia o nível socioeconômico médio da escola, isso tanto entre as escolas municipais como as estaduais (Gráfico 2).

Gráfico 2 – Distribuição da proficiência dos estudantes em Língua Portuguesa e em Matemática, por nível socioeconômico médio da escola e dependência administrativa das escolas localizadas no campo, Brasil – (2017).



Fonte: INEP/SAEB (2017). Elaboração do gráfico: própria.

Ainda sobre as variáveis relativas à escola, é necessário analisar se existia correlação entre as variáveis de proficiência dos estudantes com o percentual de adequação da formação docente e com a taxa de participação dos estudantes da escola na avaliação. O coeficiente de correlação entre a proficiência em Língua Portuguesa e Matemática dos estudantes era de 0,63, indicando que o bom desempenho em uma disciplina era positivamente correlacionado com o desempenho na outra. Já a adequação da formação dos docentes possuía correlação positiva moderada com a proficiência dos estudantes em Língua Portuguesa e Matemática: 0,17 e 0,18, respectivamente. No caso da taxa de participação da escola na avaliação, o coeficiente de correlação era fraco e positivo: 0,12 com Língua Portuguesa e 0,14 com Matemática.

É importante analisar também a variação da proficiência média dos estudantes em Língua Portuguesa e em Matemática a partir de algumas características individuais/familiares que são inclusas entre as variáveis contextuais do estudante. Isso envolve desde fatores relativos ao incentivo que os pais dão para os estudos, o trabalho fora de casa, reprovação, posse de bens e escolaridade dos pais (Tabela 2).

Entre as variáveis investigadas chama a atenção, inicialmente, a diferença entre os estudantes que foram reprovados e os que não foram. Como exemplo, os que já tinham sido reprovados ao menos uma vez apresentavam proficiência média em Língua Portuguesa de 220 pontos enquanto os que não tinham sido reprovados apresentavam média de 248,7. Situação

parecida era observada em Matemática, em que os valores eram de 221,7 e 247,6, respectivamente (Tabela 2).

O trabalho infantil e juvenil, apesar da redução, ainda constitui fenômeno histórico que caracteriza parcela das áreas rurais brasileiras. A comparação entre os que precisavam trabalhar fora com os que não precisavam apontam diferenças nos resultados. Ele implica a impossibilidade de uma dedicação mais integral aos estudos, desgaste físico e psicológico, entre outros aspectos que podem acarretar prejuízos de aprendizagem. Na amostra de estudantes residentes no campo que foi analisada, as diferenças pontuais entre os que precisam trabalhar fora de casa e aqueles que não precisavam se verificavam tanto em Língua Portuguesa como em Matemática, sugerindo a relevância desse fator social e econômico para compreender o desempenho dos estudantes (Tabela 2).

Os estudantes que tinham pais com nível médio ou superior apresentavam, em média, desempenho maior, tanto em LP como em MT. Isso sugere, em coerência com a literatura sobre fatores associados, que a escolaridade dos pais também exibiu uma correlação positiva com o desempenho escolar entre os estudantes em escolas localizadas no campo, como visto em outros grupos: “vale destacar que a escolaridade dos pais, apesar de aparecer com efeito positivo em somente três estudos, em geral tem apresentado uma relação forte e consistente, sobretudo quando se trata da escolaridade da mãe” (KARINO; LAROS, 2017, p. 114).

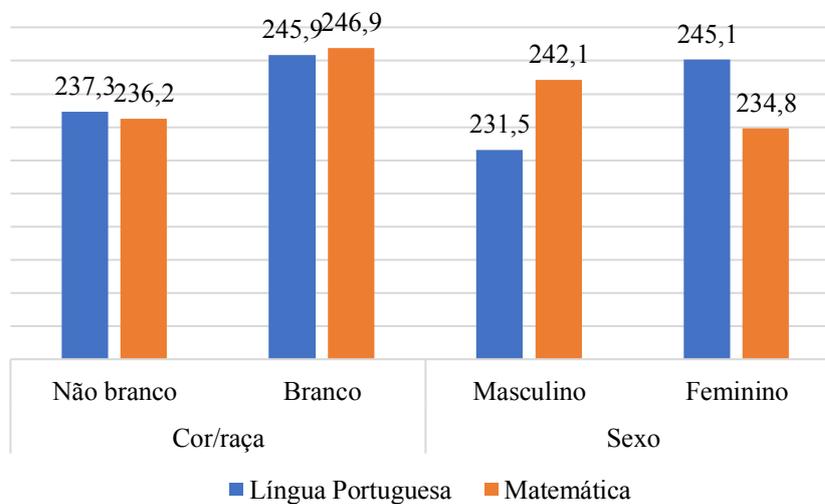
Tabela 2 – Média de proficiência dos estudantes em Língua Portuguesa e Matemática para o 9º ano do ensino fundamental em escolas localizada no campo, por práticas familiares, de estudo, escolaridade dos pais e posse de bens – Brasil (2017).

Item	Língua Portuguesa		Matemática	
	Não	Sim	Não	Sim
Seus pais incentivam você a estudar	208,1	239,0	214,4	238,8
Seus pais incentivam você a ler	236,6	238,7	239,6	238,3
Você frequenta a biblioteca?	241,5	238,7	241,2	238,7
Você lê livros?	231,8	241,6	238,5	240,9
Você trabalha (remunerado) fora de casa?	240,7	223,8	239,2	232,1
Você já foi reprovado?	248,7	220,0	247,6	221,7
Gosta de estudar Língua Portuguesa?	236,4	240,6	247,6	237,7
Gosta de estudar Matemática?	236,9	240,9	227,2	244,5
Tem carro em casa?	233,3	247,2	232,0	249,6
Tem computador em casa?	233,3	250,6	232,6	252,1
Mãe tem ensino médio ou superior?	235,4	250,3	235,3	249,9
Pai tem ensino médio ou superior?	235,7	250,8	236,5	250,1

Fonte: INEP/SAEB (2017). Elaboração própria.

Como variáveis de controle também foram inseridas informações sobre o sexo e a raça/cor (brancos e não brancos) dos estudantes das escolas localizadas nas áreas rurais (Gráfico 3). As diferenças pontuais mais amplas nas variáveis dependentes proficiência em Língua Portuguesa e Matemática se verificavam na análise por sexo. Os valores pontuais não implicavam diferenças relativas à posição nas respectivas escalas, ficando a média em Língua Portuguesa dos estudantes do sexo masculino (231,5) e feminino (245,1) alocados dentro do nível 2 da referida escala, mas aponta para desigualdades. O mesmo ocorria no caso do desempenho em Matemática, dado que tanto os estudantes do sexo masculino como feminino tinham desempenhos médios que se enquadravam no mesmo nível da escala, apesar do valor pontual relativo aos meninos ser maior (Gráfico 3). No caso da comparação por raça/cor optou-se por comparar brancos com os não-brancos, que inclui pretos, pardos, amarelos e indígenas (Gráfico 3). A análise dos valores pontuais nas proficiências registrou diferenças pequenas e que serão analisadas nos modelos de regressão.

Gráfico 3 – Média de proficiência dos estudantes em Língua Portuguesa e Matemática para o 9º ano do ensino fundamental em escolas localizada no campo, por cor/raça e sexo – Brasil (2017).



Fonte: INEP/SAEB (2017). Elaboração do gráfico: própria.

Feita a análise descritiva dos dados da amostra de estudantes do 9º ano do ensino fundamental em escolas públicas (estaduais e municipais) localizadas no campo, o trabalho passa agora a considerar os modelos de regressão linear múltipla e os modelos multinível, com a intenção de entender quais fatores associados apresentam maior relevância para a compreensão do desempenho dos estudantes e as características do efeito-escola nas áreas rurais.

2.4 Os modelos de regressão e as variáveis utilizadas

O trabalho desenvolve, inicialmente, modelos de regressão linear múltipla e, na sequência, modelos de regressão linear multinível. Nos modelos de regressão múltipla a intenção foi incluir as variáveis independentes que sejam expressivas dos fatores contextuais da escola e do estudante, de modo a compreender a associação desses com o desempenho. “A análise de regressão múltipla é mais receptiva à análise *ceteris paribus*, pois ela nos permite controlar *explicitamente* muitos outros fatores que, de maneira simultânea, afetam a variável dependente” (WOOLDRIDGE, 2016, p. 64, grifos do autor). Os modelos de regressão múltipla são importantes para que se compreenda a associação entre variáveis que, no caso em investigação, trata-se de um conjunto de dados em corte transversal do ano de 2017.

Antes de trazer os resultados é importante descrever em detalhes as variáveis utilizadas como dependentes e independentes nos modelos de regressão linear múltipla. Relacionadas à escola foram introduzidas: a variável relativa ao percentual de professores que tem formação adequada; variáveis de infraestrutura, como as relativas à escola ter *internet* e biblioteca, que oferecem uma *proxy* sobre as condições para o desenvolvimento do trabalho pedagógico; a taxa de participação dos estudantes da referida série; o Nível Socioeconômico médio da escola (relativo ao ano de 2015), que oferece uma aproximação acerca do perfil de estudantes atendidos em cada estabelecimento de ensino; e a dependência administrativa da escola, para fins de controle (Quadro 1).

Já em relação ao estudante foram introduzidas variáveis indicativas do acesso a bens de consumo diferenciadores como a posse de carro e computador; a escolaridade dos pais, comparando os que tinham nível médio e/ou superior completo aos demais; se os pais incentivam a estudar; se o estudante gosta de ler livros; se trabalha fora; se já foi reprovado; bem como a proficiência em Matemática ou Língua Portuguesa, conforme a variável dependente em análise; também foram incluídas as variáveis de sexo e cor/raça para fins de controle (Quadro 1).

Quadro 1 – Variáveis do SAEB utilizadas nos modelos de regressão múltipla

Variáveis dependentes	Grupo	Variáveis independentes
Proficiência em LP Proficiência em MT	Escola	<ul style="list-style-type: none"> - Percentual de adequação da formação Docente; - Escola tem biblioteca? (0: não; 1 sim); - Taxa de participação na prova (contínua); - NSE da escola (0 a 5); - Dependência administrativa da escola (0: estadual, 1: municipal)
	Estudante	<ul style="list-style-type: none"> - Tem carro em sua casa? (0: não; 1 sim) - Tem computador em casa? (0: não; 1 sim) - Trabalha empregada na sua casa? (0: não; 1 sim) - Mãe tem ensino médio ou superior? (0: não; 1 sim); - Pai tem ensino médio ou superior? (0: não; 1 sim) - Pais incentivam estudar? (0: não; 1 sim) - Você lê livros? (0: não; 1 sim) - Você trabalha fora de casa? (0: não; 1 sim) - Você já foi reprovado? (0: não; 1 sim) - Você gosta e estudar LP ou MT? (0: não; 1 sim); - Proficiência em Matemática/Língua portuguesa; - Sexo (0 masculino; 1 feminino) - Cor (0 não-brancos; 1 brancos)

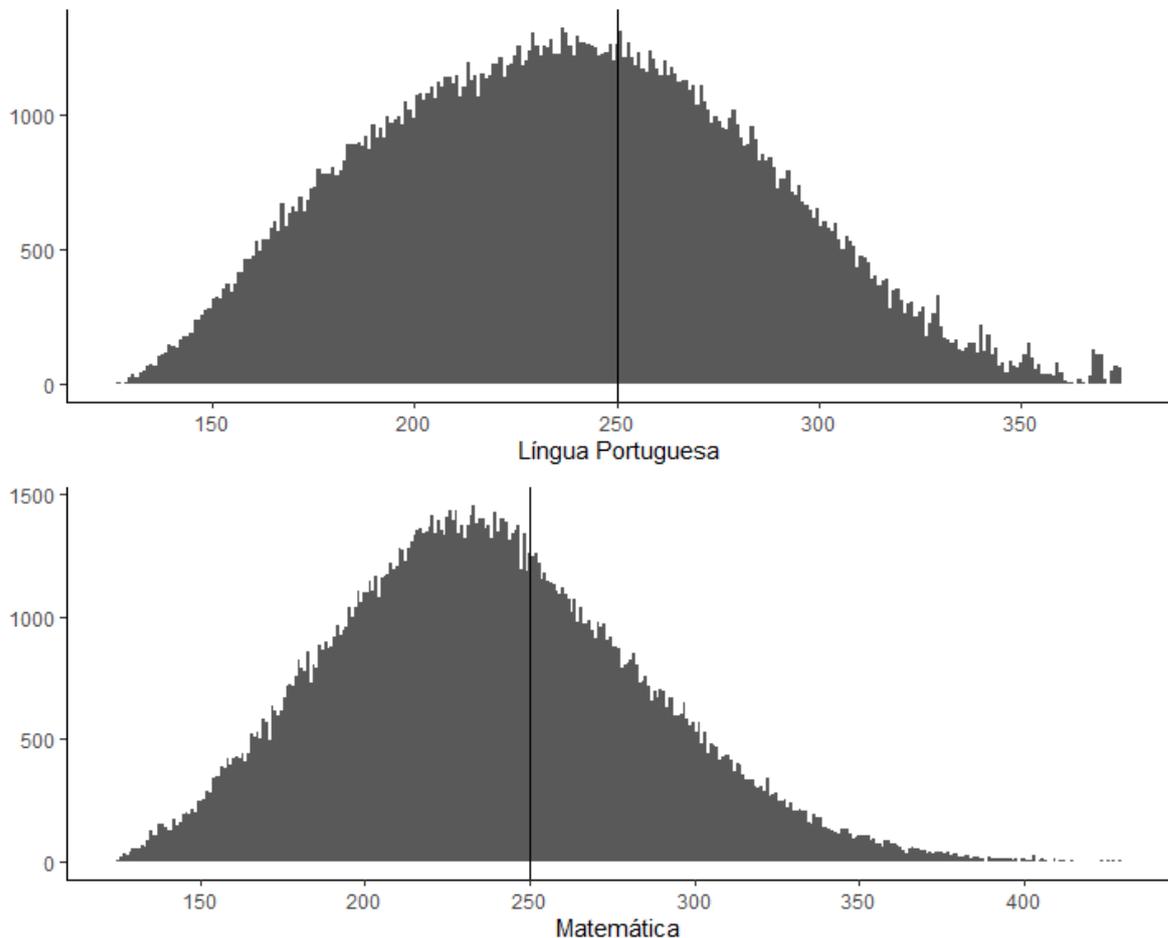
Fonte: INEP/SAEB (2018). Elaboração do quadro: própria.

2.5 Análise dos modelos de regressão linear múltipla

Inicialmente, foram ajustados modelos de regressão linear múltipla tendo como variável dependente a proficiência em Matemática e a na disciplina de Língua Portuguesa. Em cada caso, foram feitas análises considerando as variáveis da escola, as variáveis do estudante e o modelo com variáveis de ambos os níveis.

O Gráfico 4 apresenta a distribuição da proficiência dos estudantes do 9º ano do ensino fundamental em escolas públicas (estaduais e municipais) localizadas nas áreas rurais, em ambas as disciplinas (LP e MT), que são utilizadas como variáveis dependentes nos modelos. Na referida amostra, Língua Portuguesa apresentava média de 237,6, mediana de 237,1 e desvio padrão de 46,7. Em Matemática a média era de 237,5, a mediana de 234,8 e o desvio padrão de 47,1.

Gráfico 4 – Distribuição da proficiência em Língua Portuguesa e Matemática para os estudantes do 9º ano do ensino fundamental em escolas localizada no campo – Brasil (2017).



Fonte: INEP/SAEB (2017). Elaboração própria.

No caso de Matemática, são apresentados três modelos de regressão: um apenas tendo como variáveis independentes as relativas à escola (modelo 1); um apenas com as variáveis caracterizadoras dos estudantes e de sua família (modelo 2) e o modelo completo, com ambos os grupos de variáveis (modelo 3). O modelo de regressão múltipla para Matemática, com todas as variáveis, apresentou um R^2 de 0,47, o que permite dizer que as variáveis inclusas no modelo eram responsáveis por explicar 47% da variação na proficiência em Matemática (Tabela 3).

Tabela 3 – Coeficientes dos modelos de regressão múltipla para a proficiência em Matemática dos estudantes do 9º ano do ensino fundamental em escolas públicas localizadas no campo como variável dependente – Brasil (2017).

	Modelo 1	Modelo 2	Modelo 3
(Intercept)	154.12 ***	91.05 ***	63.18 ***
	(1.14)	(1.21)	(1.90)
Adequação formação docente	0.16 ***		0.05 ***
	(0.01)		(0.01)
Taxa de participação	57.20 ***		25.58 ***
	(1.15)		(1.37)
Tem Internet?	4.40 ***		1.54 ***
	(0.26)		(0.30)
Tem biblioteca?	5.51 ***		2.37 ***
	(0.27)		(0.32)
NSE escola	8.50 ***		2.61 ***
	(0.17)		(0.21)
Dependência administrativa	1.35 ***		1.40 ***
	(0.35)		(0.39)
Proficiência em LP		0.62 ***	0.60 ***
		(0.00)	(0.00)
Pais incentivam estudar		1.47	1.16
		(0.94)	(1.13)
Lê livros		-2.87 ***	-2.59 ***
		(0.44)	(0.51)
Trabalha fora		-0.15	-0.67
		(0.37)	(0.44)
Foi reprovado		-5.84 ***	-5.45 ***
		(0.20)	(0.24)
Gosta de estudar matemática		13.84 ***	14.61 ***
		(0.29)	(0.34)
Tem carro em casa		6.17 ***	4.46 ***
		(0.28)	(0.34)
Tem computador em casa		4.92 ***	3.33 ***
		(0.30)	(0.36)
Mãe tem médio ou superior		0.79 *	0.88 *
		(0.32)	(0.38)
Pai tem médio ou superior		-0.35	-0.44
		(0.36)	(0.42)
Raça/cor		3.34 ***	1.82 ***
		(0.29)	(0.35)
Sexo		-16.38 ***	-16.55 ***
		(0.26)	(0.31)
N	126177	76993	54128
R2	0.08	0.47	0.47

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Fonte: INEP/SAEB (2017). Elaboração própria. Para informações sobre ajuste do modelo ver Anexo A.

Na análise do modelo 3, em que todas as variáveis foram incluídas, é possível notar que, entre as relativas à escola, a ampliação em um ponto na taxa de participação – que varia, na amostra, entre de 0,08 a 1,2 – implicava um aumento, em média, de 25,58 pontos na proficiência em Matemática. A existência de *internet* e biblioteca na escola também apresentavam coeficientes positivos, bem como o nível socioeconômico médio. Já em relação às variáveis que podem ser definidas como associadas às condições dos estudantes e da família, posse de bens como carro e computador em casa, por exemplo, registravam coeficientes positivos, assim como o fato de o estudante gostar de estudar Matemática, que apresentava um desempenho de 14,41 pontos maiores do que os que afirmavam não gostar. Já a situação de o estudante ter sido reprovado apresentava um coeficiente negativo: os que tinham reprovado apresentavam um desempenho, em média, 5,45 pontos inferiores aos que não tinham. Cabe destacar também que as estudantes do sexo feminino apresentavam, em média, considerando a amostra e as variáveis incluídas no modelo, um desempenho em Matemática de 16,55 pontos menores que o obtido pelos estudantes do sexo masculino (Tabela 3). Da análise do modelo, nota-se que os fatores socioeconômicos possuíam relevância explicativa no desempenho de Matemática, em coerência com aspectos apontados pela literatura (BROOKE; SOARES, 2008).

No caso de Língua Portuguesa como variável dependente – que apresenta três modelos, tal como o de Matemática – a taxa de participação apresentou coeficiente positivo e estatisticamente significativo no modelo 3 (completo). O mesmo pode ser dito em relação à existência de equipamentos de infraestrutura: estudantes de escolas que tinham *internet* apresentavam um desempenho de 1,29 pontos, em média, maiores aos que estudavam em escolas que não possuíam; situação similar podia ser vislumbrada no caso das escolas que possuíam biblioteca. A adequação da formação docente, apresentava coeficiente positivo, mas baixo. No que se refere às variáveis relativas ao estudante, os que afirmavam que os pais incentivavam a estudar apresentavam desempenho, em média, de 10,64 pontos superiores aos que afirmavam que não incentivavam. Também eram positivos e estatisticamente significativos os coeficientes relativos a gostar de estudar Língua Portuguesa e de ler livros. As estudantes do sexo feminino apresentavam um desempenho em Língua Portuguesa que era, em média, 15,96 pontos superiores aos do sexo masculino. Importante destacar que essa diferença era inversa ao registrado no desempenho em Matemática, no qual o desempenho dos meninos era maior, fenômeno coerente com os achados da literatura (KARINO; LAROS, 2017). Já os estudantes que tinham que trabalhar fora apresentavam um desempenho que era 5,13 pontos inferiores aos que não precisavam, em média; situação similar se constatava em

relação aos que haviam sido reprovados: desempenho que era 6,02 pontos inferiores aos que não tinham sido (Tabela 4).

Tabela 4 – Coeficientes dos modelos de regressão múltipla para a proficiência em Língua Portuguesa dos estudantes do 9º ano do ensino fundamental em escolas públicas localizadas no campo como variável dependente – Brasil (2017).

	Modelo 1	Modelo 2	Modelo 3
(Intercept)	161.87 *** (1.14)	70.74 *** (1.23)	53.46 *** (1.91)
Adequação formação docente	0.15 *** (0.01)		0.06 *** (0.01)
Taxa de participação	50.85 *** (1.15)		13.82 *** (1.37)
Internet	4.03 *** (0.26)		1.29 *** (0.30)
Biblioteca	4.76 *** (0.27)		1.32 *** (0.32)
NSE escola	8.07 *** (0.17)		2.15 *** (0.21)
Dependência administrativa	1.30 *** (0.35)		0.03 (0.39)
Proficiência em Matemática		0.60 *** (0.00)	0.58 *** (0.00)
Pais incentivam estudar		10.63 *** (0.94)	10.64 *** (1.12)
Lê livros		2.42 *** (0.44)	3.29 *** (0.52)
Trabalha fora		-5.51 *** (0.36)	-5.13 *** (0.43)
Foi reprovado		-6.10 *** (0.20)	-6.02 *** (0.24)
Gosta de estudar Língua Portuguesa		7.04 *** (0.33)	7.65 *** (0.39)
Tem carro em casa		1.56 *** (0.28)	-0.49 (0.34)
Tem computador em casa		3.78 *** (0.30)	2.91 *** (0.36)
Mãe tem médio ou superior		2.71 *** (0.32)	2.65 *** (0.37)
Pai tem médio ou superior		3.03 *** (0.35)	2.61 *** (0.42)
Raça/cor		1.83 *** (0.29)	1.05 ** (0.35)
Sexo		15.88 *** (0.26)	15.96 *** (0.31)
N	126177	77193	54300
R2	0.07	0.46	0.46

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Fonte: INEP/SAEB (2017). Elaboração própria. Para informações sobre ajuste do modelo ver Anexo B.

A comparação entre os coeficientes encontrados em Matemática e em Língua Portuguesa sugere diferenças quanto aos fatores associados. No caso de LP, o fato de ter de trabalhar fora apresentava coeficientes superiores aos encontrados no caso de Matemática. O mesmo ocorria com o atraso escolar, medido pela variável indicadora de reprovação. Isso, novamente, é coerente com a literatura sobre fatores associados, que tem apontado “que, quando o estudante é reprovado em algum momento, ele tende a ter um desempenho inferior aos demais estudantes no decorrer da trajetória escolar, mesmo tendo cursado algumas vezes uma determinada série” (KARINO; LAROS, 2017, p. 114).

Também em relação à escolaridade dos pais, em LP o fato de a mãe e/ou o pai terem ensino médio ou superior estava associado com desempenhos maiores, já em Matemática não é possível, na amostra considerada, encontrar tal associação. No que se refere às características da escola, os coeficientes foram positivos e significativos, tanto para Língua Portuguesa como para Matemática (Tabela 3; Tabela 4).

Como discutido anteriormente, os modelos de regressão linear múltipla são importantes para uma apreensão acerca das associações que podem existir entre fatores escolares e extraescolares. Eles possuem algumas limitações, especialmente relacionadas ao pressuposto de independência das observações, dado que os estudantes estudam em escolas diferentes e estando dentro de uma mesma instituição compartilham características que limitam o pressuposto de independência entre as observações. Isso pode ser ainda mais relevante nas áreas rurais, onde as possibilidades de selecionar a escola são mais restritas, seja em função das distâncias e da própria existência de um número reduzido de estabelecimentos por comunidade que, muitas vezes, contam com apenas uma instituição de ensino.

Na sequência são analisados os modelos multiníveis de modo a subsidiar a reflexão sobre os fatores associados ao desempenho nas escolas localizadas no campo e acerca das peculiaridades do efeito-escola em tais espaços.

2.6 Análise dos modelos de regressão multinível

Anteriormente, nos modelos de regressão linear múltipla, foram utilizadas variáveis relativas à escola e ao estudante. Uma questão teórica que surge, contudo, é que as características dos dados sobre a escola pressupõem uma forma de hierarquização em níveis distintos. A interação dos estudantes ocorre no ambiente comunitário e familiar, mas também no interior da escola, ou mesmo em uma turma, que agrega estudantes distintos em um mesmo estabelecimento. Cada espaço de interação pode acarretar diferenças nos resultados,

dado que eles contêm características que implicam influências distintas sobre o aprendizado, as competências e habilidades que podem afetar o desempenho dos estudantes nas avaliações.

Os dados coletados nas áreas sociais e humanas são frequentemente de pessoas agrupadas em *clusters*, também conhecidos como conglomerados (estudantes em escolas, indivíduos em sociedades, trabalhadores em empresas); logo, é provável que compartilhem atributos similares em decorrência do contexto que lhes é comum. Como consequência da dependência das observações ou atributos mensurados, ocorre a subestimação dos erros-padrão dos coeficientes da regressão. Uma alternativa para a regressão múltipla, que leva em consideração essa similaridade, é a análise multinível. Esta técnica é um tipo de análise de regressão que contempla simultaneamente múltiplos níveis de agregação, tornando corretos os erros-padrão, os intervalos de confiança e os testes de hipóteses (PUENTE-PALACIOS; LAROS, 2009, p. 350).

Na atualidade, as investigações sobre o efeito-escola têm empregado com frequência modelos multinível, em função das características dos dados educacionais (FERRÃO, 2003; ANDRADE; LAROS, 2007, PUENTE-PALACIOS; LAROS, 2009). Entre as vantagens dos modelos multiníveis está apreender os “efeitos combinados tanto das variáveis individuais como das de grupos” (CRUZ, 2010, p. 8). No caso dos modelos de regressão linear multinível aqui empregados foi feita uma análise em dois níveis: estudante (nível 1) e escola (nível 2). A técnica auxilia em apreender o efeito da escola sobre a proficiência, distinguindo-o do que pode ser proveniente das características do estudante.

O ponto de partida para uma análise multinível é entender se a variação nos dados efetivamente pode estar relacionada ao caráter hierárquico das informações, daí a importância de partir do modelo nulo, em que apenas a variável independente de interesse e a indicadora do nível (no caso o código da escola) estejam presentes (SNIJDERS; BOSKER, 2012; HECK; THOMAS; TABATA, 2010; FIELD, 2020; FERRÃO, 2003).

Para fazer a interpretação dos valores é comumente utilizado o Coeficiente de Correlação Intraclasse (CCI), que “é um bom indicador da extensão em que uma variável contextual tem um efeito sobre o resultado” (FIELD, 2020, p. 939). Ele representa o percentual da variabilidade total que é dependente da variável de nível 2, no caso em análise, a escola. Como explica Ferrão (2003), tratando do seu uso na análise educacional, “o coeficiente de correlação intra-escolar (...) mede a proporção da variância entre escolas face a variância total da variável resposta. Esta estatística permite ao investigador ter melhor noção da magnitude do efeito-escola” (p. 34). Ele resulta de uma razão entre a variância entre escolas sobre a soma da variância total.

O coeficiente varia de 0 a 1. Quando seu valor é nulo, significa que as escolas são homogêneas entre si e que o desempenho escolar do aluno independe da escola que ele frequenta, ou seja, toda variabilidade dos resultados escolares deve-se apenas às diferenças entre os alunos. Na situação extrema, quando o coeficiente de correlação intra-escolar tem valor 1, toda a variabilidade no desempenho dos alunos deve-se à diferença entre as escolas e, em tal situação hipotética, as características individuais do aluno em nada afetam o seu desempenho escolar, ficando este a dever-se inteiramente às características da escola que ele frequenta (FERRÃO, 2003, p. 34).

Assim, por um lado, quanto mais o desempenho dos estudantes dentro de uma escola for homogêneo entre eles e diferente do observado nas demais escolas, maior o efeito da escola, dado que aquele estabelecimento em particular pode ter efeito significativo sobre o aprendizado dos seus estudantes. Por outro lado, quanto mais variedade ocorrer entre os estudantes como um todo e menos diferentes eles forem entre as escolas, menos efeitos essas têm.

O primeiro modelo nulo analisado foi com a proficiência em Matemática. Existiriam diferenças entre os desempenhos em MT dos estudantes em escolas no campo ao se considerar o nível da escola? No caso da proficiência em Matemática, a média global (γ_{00}) foi de 236,8. Tanto a variação entre (*between*) as escolas (*intercept* = 608,6, *Wald Z* = 55,7, $p < 0,001$), como dentro do grupo dos estudantes (*residuals* = 1629,3, *Wald Z* = 293,2, $p < 0,001$) foram significativas. Já o CCI observado na amostra foi de 0,272, o que permite afirmar que, no caso de Matemática, 27,1% da variação no desempenho dos estudantes está relacionada às escolas. Em outras palavras existem diferenças suficientes no desempenho dos estudantes entre as escolas.

O segundo modelo nulo foi o de Língua Portuguesa. A média global (γ_{00}) foi de 236,7, a variação entre as escolas foi estatisticamente significativa (*intercept* = 514,0, *Wald Z* = 53,9, $p < 0,001$), assim como dentro dos estudantes (*residuals* = 1691,9, *Wald Z* = 293,2, $p < 0,001$). Já o CCI foi de 0,233, valor que indica a existência de uma variação de 23,3% entre as escolas sobre o desempenho dos estudantes no contexto rural, indicando validade de se considerar a estrutura de agrupamento na investigação.

Após a análise dos modelos nulos, as análises utilizaram como variáveis independentes da escola (nível 2): o indicador de adequação da formação docente (variável contínua), a taxa de participação na prova (variável contínua), se existe *internet* na escola (0: não; 1: sim) e o nível socioeconômico médio da escola (variável ordinal de 1 a 5, no caso da amostra), que, apesar de ser composta a partir de características socioeconômicas e familiares

dos estudantes de um estabelecimento de ensino, representa “o padrão de vida de seu público, relacionados à respectiva posição na hierarquia social” (INEP, 2015, sem página). Em relação ao nível 1, do estudante, foram utilizadas: a proficiência em LP ou MT, se os pais incentivam estudar (0: não; 1: sim), se o estudante trabalhava fora (0: não; 1: sim), sobre ter sido ou não reprovado (0: não; 1: sim), se gosta de estudar LP ou MT (0 e 1), se a mãe tem ensino médio completo ou superior (0: não; 1: sim), cor/raça (não-brancos: 0; brancos: 1; nesse caso, como dito antes, optou-se por comparar os brancos aos demais grupos, agregados na categoria não-brancos) e o sexo do estudante (0: masculino; 1: feminino). Em relação às variáveis explicativas foram inclusas como efeito aleatório, no nível da escola, a taxa de participação e o NSE médio da escola; já no nível do aluno foram inclusos os efeitos aleatórios para variável referente ao fato do estudante ter sido ou não reprovado e se ele trabalhava⁸. Além disso, o modelo utiliza o intercepto aleatório da escola.

Os parâmetros fixos do modelo, após os ajustes citados anteriormente, apresentaram informações significativas. Destaque para a taxa de participação, para a qual a ampliação de um ponto implicava o aumento de, em médias 23,7 pontos no desempenho de Matemática; no nível do aluno o fato de trabalhar fora implicava nota de -0,8 menor que os que não trabalhavam. Chama atenção como o fato de ter sido reprovado implicava uma média de 9,26 pontos a menos frente aos que não haviam sido reprovados (Tabela 5). Por fim, cabe ressaltar que mesmo com os controles inseridos, bem como a modelagem aleatória das inclinações, as estudantes do sexo feminino apresentavam desempenho em Matemática, em média, de -16,49 pontos a menos que os estudantes no sexo masculino (Tabela 5). Interessante notar também que, mesmo com a inserção das citadas variáveis, a análise da variação entre e intra-escola apontava para a permanência de um efeito das escolas nas áreas rurais em Matemática, considerando a amostra que foi utilizada (Anexo C).

Tabela 5 – Coeficientes para o modelo multinível tendo Matemática como variável dependente para os estudantes do 9º ano do ensino fundamental em escolas no campo – Brasil (2017).

Parâmetros fixos	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	75,063	1,821	146994,096	41,228	0,000	71,494	78,631
NSE escola	3,523	0,271	361577,709	13,009	0,000	2,992	4,053
Adequação formação	0,074	0,009	357188,029	8,065	0,000	0,056	0,093
Taxa de participação	23,760	1,602	219564,343	14,827	0,000	20,619	26,900
Internet na escola	2,336	0,405	372787,306	5,767	0,000	1,542	3,129
Proficiência em LP	0,548	0,003	81944,910	186,418	0,000	0,543	0,554

⁸ O modelo foi realizado no SPSS/PASW (18). Utilizou-se os componentes de variância.

Pais incentivam estudar	1,706	0,881	76724,250	1,937	0,053	-0,021	3,433
Trabalha fora	-0,803	0,377	3073,804	-2,131	0,033	-1,541	-0,064
Reprovação	-9,264	0,282	86604,813	-32,907	0,000	-9,815	-8,712
Gosta de MT	14,155	0,275	78573,253	51,424	0,000	13,615	14,694
Estudo da mãe	2,874	0,279	78845,916	10,282	0,000	2,326	3,422
Raça/cor	1,103	0,296	80993,690	3,730	0,000	0,523	1,682
Sexo	-16,490	0,250	77009,086	-65,966	0,000	-16,980	-16,000

Fonte: INEP/SAEB (2018). Elaboração do modelo: própria.

Nota: parâmetros aleatórios e saídas do modelo nulo disponíveis no Anexo C.

No caso do modelo multinível para Língua Portuguesa, todos os parâmetros fixos apresentavam coeficientes estatisticamente significativos (p -valor $< 0,001$). No nível da escola, a taxa de participação na prova e a existência de *internet* na escola, por exemplo, apresentavam associação positiva e estatisticamente significativa com o desempenho em Língua Portuguesa. Novamente, chama a atenção que, mesmo com a inclusão dos coeficientes aleatórios e dos controles, no nível do estudante, o fato de ter reprovado implicava uma nota, em média, 9,37 inferiores, assim como trabalhar fora, para os quais a nota média era 5,41 menor do que a dos que não trabalhavam. O fato de a mãe possuir nível médio ou superior estava correlacionado com um desempenho de 4,29 pontos maior em relação aos demais (Tabela 6). Diferente do observado para Matemática, no qual o desempenho das estudantes do sexo feminino era menor, no caso do modelo que tem Língua Portuguesa como variável dependente, considerando os estudantes do 9º ano do ensino fundamental em escolas públicas localizadas no campo, a nota das estudantes do sexo feminino era 15,03 pontos superior a nota dos estudantes do sexo masculino (Tabela 6). Ressalta-se que, mesmo com a importância explicativa trazida pelos fatores contextuais, a variação entre os estabelecimentos ainda seguia indicando a existência de efeitos da escola que precisam ser melhor analisados nas áreas rurais (Anexo D).

Tabela 6 – Coeficientes para o modelo multinível tendo Língua Portuguesa como variável dependente para os estudantes do 9º ano do ensino fundamental em escolas no campo – Brasil (2017).

Parâmetros fixos	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	60,824	1,808	11573,291	33,646	0,000	57,280	64,367
NSE escola	2,901	0,233	4482,997	12,454	0,000	2,444	3,358
Adequação formação	0,063	0,008	4225,374	7,750	0,000	0,047	0,079
Taxa de participação	14,842	1,529	5331,383	9,707	0,000	11,844	17,839
Internet na escola	1,713	0,355	4086,455	4,823	0,000	1,017	2,409
Proficiência em MT	0,564	0,003	69231,624	189,396	0,000	0,558	0,570
Pais incentivam estudar	10,395	0,903	74996,734	11,510	0,000	8,625	12,165
Trabalha fora	-5,414	0,374	3013,498	-14,492	0,000	-6,146	-4,681
Reprovação	-9,372	0,284	5249,603	-33,009	0,000	-9,928	-8,815

Gosta de LP	7,809	0,331	76178,560	23,607	0,000	7,161	8,458
Estudo da mãe	4,297	0,284	75989,713	15,107	0,000	3,739	4,854
Raça/cor	0,964	0,300	74294,902	3,210	0,001	0,375	1,552
Sexo	15,306	0,258	75780,751	59,262	0,000	14,799	15,812

Fonte: INEP/SAEB (2018). Elaboração do modelo: própria.

Nota: parâmetros aleatórios e saídas do modelo nulo disponíveis no Anexo D.

3 CONCLUSÕES

As avaliações externas em larga escala, como é o caso do SAEB, oferecem recursos importantes para que se compreenda aspectos da qualidade educacional relacionados ao aprendizado dos estudantes, em diversos contextos. A análise do desempenho em Língua Portuguesa e em Matemática dos estudantes do 9º ano em escolas localizadas no campo permitiu aprofundar tal compreensão.

O acesso e a qualidade do ensino nas áreas rurais constituem desafios históricos às políticas educacionais e tem gerado algumas ações visando melhorar as condições escolares. Frente a isso, foi importante compreender como aspectos relacionados à infraestrutura escolar se correlacionam com o aprendizado, tanto em Língua Portuguesa como em Matemática, o que se mostrou coerente com achados da literatura sobre fatores associados (KARINO; LAROS, 2017). Ao mesmo tempo, ainda com o uso de modelos de regressão linear múltipla, foi possível incluir variáveis que apontaram para a associação de aspectos como o nível socioeconômico, a escolaridade dos pais, a necessidade de trabalhar fora e o atraso escolar com o desempenho dos estudantes nas áreas rurais, como registrado em outros espaços.

Com os modelos de regressão multinível foi possível observar que existe um efeito positivo da taxa de participação da escola e de itens de infraestrutura, como a existência de *internet*, por exemplo. No nível do estudante, nota-se que a escolaridade da mãe e o fato de gostar da disciplina em análise apresentavam efeitos positivo sobre o desempenho. Todavia, destaca-se que, mesmo com a introdução dos controles, trabalhar fora e ter sido reprovado eram situações que apresentavam maior associação negativa com o desempenho entre os estudantes em escolas nas áreas rurais. A comparação entre os modelos nulos e os modelos com variáveis explicativas inclusas sugere a existência de um efeito-escola nas áreas rurais dentro da amostra selecionada, que precisa ser compreendido em suas especificidades.

O estudo trouxe alguns subsídios para entender a relevância das avaliações na compreensão da qualidade educacional nas áreas rurais. Também permitiu reunir informações para entender os desafios à garantia do aprendizado nos distintos territórios e espaços em que a escola se insere, considerando a importância dos contextos. É preciso avançar na

compreensão das especificidades dos fatores contextuais e do efeito-escola no campo, a partir de suas diversidades e de suas transformações ao longo das últimas décadas.

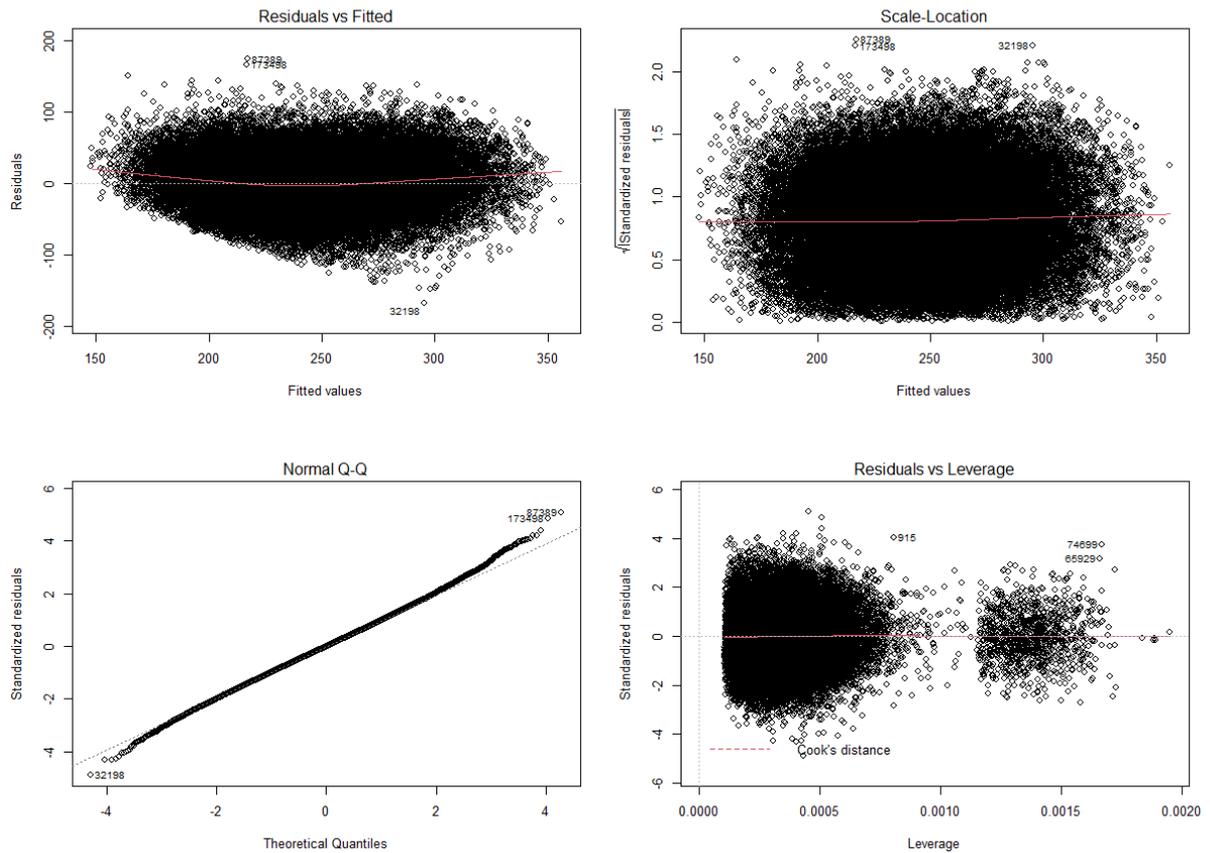
4 REFERÊNCIAS

- ARROYO, Miguel Gonzales; CALDART, Roseli Salet; MOLINA, Mônica Castagna. *Apresentação*. In: ARROYO, Miguel Gonzales; CALDART, Roseli Salet; MOLINA, Mônica Castagna (orgs.). *Por uma educação do campo*. 5ª edição. Petrópolis, RJ: Vozes, 2011.
- ALVES, Maria Teresa Gonzaga; SOARES, José Francisco. *Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos alunos*. *Educ. rev.*, Belo Horizonte, n. 45, p. 25-59, Junho de 2007. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-46982007000100003&lng=en&nrm=iso. Acesso em 18 de janeiro de 2021.
- ANDRADE, Josemberg M. de; LAROS, Jacob A.. *Fatores Associados ao Desempenho Escolar: Estudo Multinível com Dados do SAEB/2001*. *Revista Psicologia: Teoria e Pesquisa*. Jan-Mar 2007, Vol. 23 n. 1, pp. 033-042. Disponível em: <https://www.scielo.br/pdf/ptp/v23n1/a05v23n1.pdf>. Acesso em: 22 de fevereiro de 2021.
- ANDRADE, Renato Júdice; SOARES, José Francisco. *O efeito da escola básica brasileira*. *Revista Estudos em Avaliação Educacional*, v. 19, n. 41, set./dez. 2008. São Paulo.
- BOF, Alvana Maria (organização). *A educação no Brasil rural*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2006.
- BONAMINO, Alicia; SOUSA, Sandra Zákia. *Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola*. *Revista Educação e Pesquisa*, São Paulo, v. 38, n. 2, p. 373-388, Junho de 2012. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-97022012000200007&lng=en&nrm=iso. Acesso em 26 de outubro de 2020.
- BOURDIEU, Pierre. *Escritos de Educação*. 5ª edição. Organização: Maria Alice Nogueira; Afrânio Catani. Petrópolis: Vozes, 2003.
- BROOKE, Nigel; SOARES, José Francisco (organizadores). *Pesquisa em eficácia escolar – origens e trajetórias*. Belo Horizonte: Editora UFMG, 2008.
- COLEMAN, James S. *Desempenho nas escolas públicas*. Em: BROOKE, Nigel; SOARES, José Francisco (organizadores). *Pesquisa em eficácia escolar – origens e trajetórias*. Belo Horizonte: Editora UFMG, 2008.
- CRUZ, Cláudia Catarina Mendes Silva da. *Modelos multi-nível: fundamentos e aplicações*. Dissertação em Matemática, Estatística e Computação. Lisboa: Universidade Aberta, 2010. Disponível em: <https://repositorioaberto.uab.pt/handle/10400.2/1729>. Acesso em 24 de fevereiro de 2021.
- FAVARETO, Arilson. *A educação nos marcos das transformações do rural contemporâneo*. *Revista Educ. Soc.*, Campinas, v. 35, n. 129, p. 1137-1163, Dec. 2014. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-73302014000401137&lng=en&nrm=iso. Acessado em 22 de setembro de 2020.
- FERRÃO, Maria Eugênia. *Introdução aos modelos de regressão multinível em educação*. Campinas: Komedi, 2003.
- FERRÃO, M. E.; FERNANDES, C. *O efeito-escola e a mudança - dá para mudar? Evidências da investigação Brasileira*. REICE. *Revista Iberoamericana Sobre Calidad, Eficacia Y Cambio En Educación*, vol. 1, nº 1, 2003. Recuperado a partir de <https://revistas.uam.es/reice/article/view/5343>. Acesso em 30 de janeiro de 2021.
- FIELD, Andy. *Descobrimo a estatística usando o SPSS*. 5ª ed. Porto Alegre: Penso, 2020.

- HECK, Ronald H.; THOMAS, SCOTT L.; TABATA, Lynn N. *Multilevel and longitudinal modeling with IBM SPSS*. New York, London: Routledge, 2010.
- INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse) – Nota Técnica. Brasília: INEP, 2015. Disponível em: <http://portal.inep.gov.br/web/guest/indicadores-educacionais>. Acesso em 22 de fevereiro de 2021.
- INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sistema de Avaliação da Educação Básica (SAEB) – microdados 2017*. Brasília: INEP, 2018. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb>. Acesso em 07 de dezembro de 2020.
- KARINO, Camila Akemi; LAROS, Jacob Aries. Eficácia escolar: uma revisão de literatura. *Revista Examen*, v. 1., n. 1, p. 95-126, 2017. Disponível em: <https://examen.emnuvens.com.br/rev/article/view/25/3>. Acessado em 10 de fevereiro de 2021.
- MOLINA, Mônica Castagna; Freitas, Helana Célia de Abreu. *Avanços e desafios na construção da educação do campo*. *Revista Em Aberto*, Brasília, v. 24, n. 85, p. 17-31, abr. 2011. Disponível em: <http://rbepold.inep.gov.br/index.php/emaberto/article/view/2483/2440>. Acesso em 17 de fevereiro de 2021.
- PICANÇO, Felícia; MORAIS, Juliana. *Estudos sobre estratificação educacional: síntese dos principais argumentos e desdobramentos*. *Revista Educação e Sociedade*, Campinas, v. 37, n. 35, p. 391-406, Junho de 2016. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-73302016000200391&lng=en&nrm=iso. Acesso em 17 de fevereiro de 2021.
- PUENTE-PALACIOS, Katia Elizabeth; LAROS, Jacob Arie. Análise multinível: contribuições para estudos sobre efeito do contexto social no comportamento individual. *Revista Estud. psicol. (Campinas)*, Campinas, v. 26, n. 3, p. 349-361, Sept. 2009. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-166X2009000300008&lng=en&nrm=iso. Acessado em 28 de setembro de 2020.
- SANTOS, Robson dos. *Relações entre escolarização, pobreza e posições de classe no campo brasileiro e suas implicações para as políticas educacionais*. 2019. xiv, 227 f., il. Tese (Doutorado em Política Social) - Universidade de Brasília, Brasília, 2019.
- SNIJDERS, Tom A. B.; BOSKER, Roel J. *Multilevel Analysis – an introduction to basic and advanced multilevel modeling*. 2ª ed. London: SAGE, 2012.
- SOARES, José Francisco; ALVES, Maria Teresa Gonzaga. *Efeitos de escolas e municípios na qualidade do ensino fundamental*. *Cadernos de Pesquisa*, São Paulo, v. 43, n. 149, p. 492-517, Agosto 2013. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-15742013000200007&lng=en&nrm=iso. Acesso em 26 de outubro de 2020.
- SOARES NETO, Joaquim José; JESUS, Gírlene Ribeiro de; KARINO, Camila Akemi; ANDRADE, Dalton Francisco de. Uma escala para medir a infraestrutura escolar. *Revista Estudos de Avaliação Educacional*. v. 24 n. 54 (2013). Disponível em: <http://publicacoes.fcc.org.br/index.php/eae/article/view/1903/1887>. Acesso em 01 de fevereiro de 2021.
- VEIGA, José Eli da. Nascimento de outra ruralidade. *Revista Estudos Avançados*, São Paulo, v. 20, n. 57, p. 333-353, Agosto de 2006. Disponível em http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142006000200023&lng=en&nrm=iso. Acesso em 30 janeiro de 2021.
- WINTER, Bodo. *Linear models and linear mixed effects models in R with linguistic applications*. 2013. arXiv preprint arXiv:1308.5499, 2019. Disponível em: <https://arxiv.org/abs/1308.5499>. Acesso em 16 de fevereiro de 2021.
- WOOLDRIDGE, Jeffrey M. *Introdução à Econometria – uma abordagem moderna*. 4ª edição. Tradução de José Antônio Ferreira. São Paulo: Cengage Learning, 2016.

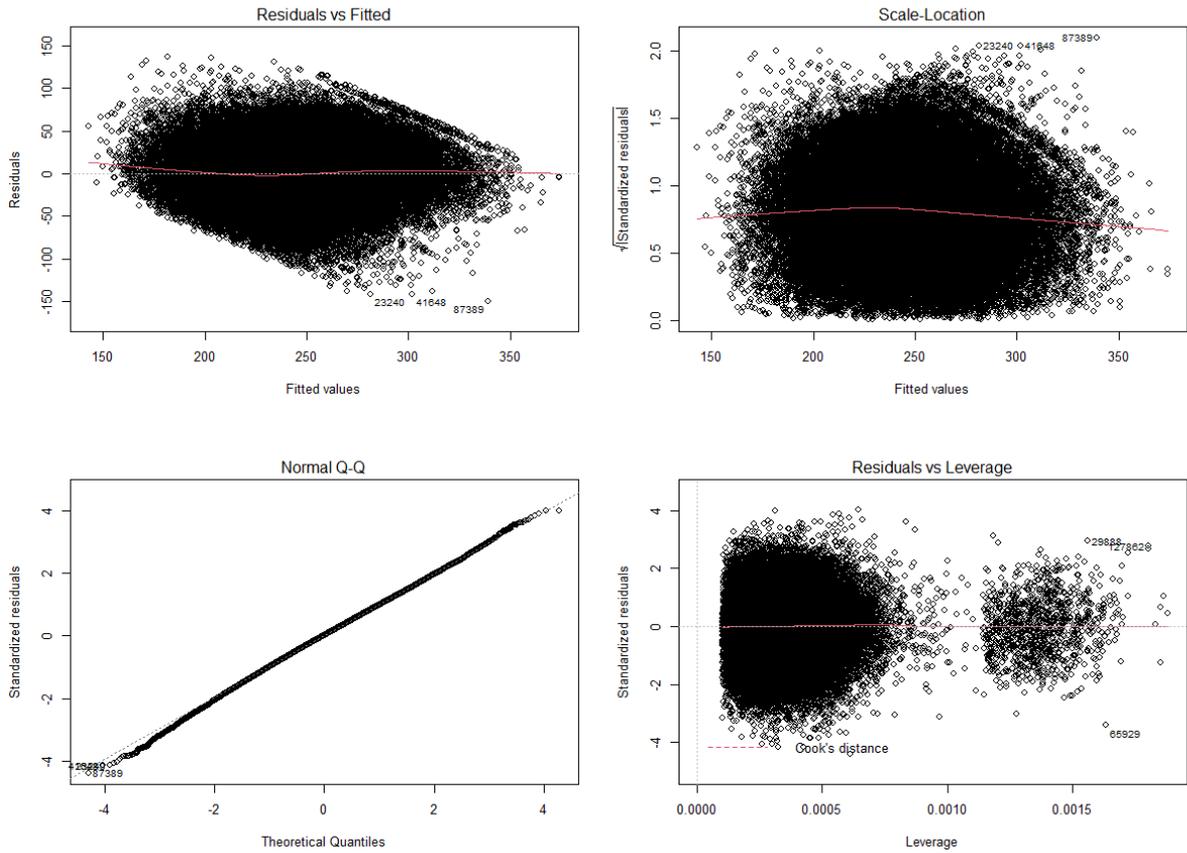
ANEXO

ANEXO A - GRÁFICOS DO AJUSTE PARA O MODELO DE REGRESSÃO LINEAR MÚLTIPLA DE MATEMÁTICA



Fonte: INEP/SAEB (2018). Elaboração do modelo: própria.
Elaboração com base no pacote R.

ANEXO B - GRÁFICOS DO AJUSTE PARA O MODELO DE REGRESSÃO LINEAR MÚLTIPLA DE LÍNGUA PORTUGUESA



Fonte: INEP/SAEB (2018). Elaboração do modelo: própria.
Elaboração com base no pacote R.

**ANEXO C – SAÍDAS MODELOS MULTINÍVEL DE MATEMÁTICA
(ELABORADOS NO SPSS/PASW - 18)**

Modelo Nulo de Matemática

Estimates of Fixed Effects ^a							95% Confidence Interval	
Parameter		Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
dimension0	Intercept	236,7976	0,290045	8320,544	816,416	0	236,229	237,3661

a. Dependent Variable: Matemática

Estimates of Covariance Parameters ^(a)							
Parameter						95% Confidence Interval	
		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		1629,29	5,56	293,25	0,00	1618,43	1640,21
Intercept [subject = ESCOLA]	Variance	608,47	10,92	55,71	0,00	587,44	630,26

a. Dependent Variable: Matemática

Parâmetros aleatórios do modelo para Matemática

Estimates of Covariance Parameters ^b							
Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		1032,784	5,461	189,134	,000	1022,137	1043,542
Intercept [subject = ESCOLA]	Variance	,038	,000
Taxa de participação [subject = ESCOLA]	Variance	89,100	,000
Nse da escola [subject = ESCOLA]	Variance	6,659	,000
Reprovação [subject = ESCOLA]	Variance	10,846	,000
Trabalha fora [subject = ESCOLA]	Variance	34,579	9,749	3,547	,000	19,899	60,088

Fonte dos dados das saídas: INEP/SAEB (2018). Elaboração do modelo: própria.

**ANEXO D – SAÍDAS MODELOS MULTINÍVEL DE LÍNGUA PORTUGUESA
(ELABORADOS NO SPSS/PASW - 18)**

Modelo Nulo de Língua Portuguesa

Parameter	Estimates of Fixed Effects(a)						95% Confidence Interval	
		Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
dimension0	Intercept	236,686	0,271	8258,991	874,492	0,000	236,155	237,216

a. Dependent Variable: Língua Portuguesa.

Parameter	Estimates of Covariance Parameters(a)					95% Confidence Interval	
		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		1691,882	5,769989	293,221	0	1680,611	1703,229
Intercept [subject = ESCOLA]	Variance	513,8932	9,530232	53,922	0	495,5497	532,9157

a. Dependent Variable: Língua Portuguesa.

Parâmetros aleatórios para o modelo de Língua Portuguesa

Estimates of Covariance Parametersb							
Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		1100,641	5,975	184,205	,000	1088,992	1112,414
Intercept [subject = ESCOLA]	Variance	62,259	2,984	20,862	,000	56,676	68,392
tx_partic_9EF_esc [subject = ESCOLA]	Variance	,000	,000
nse_recod_escola [subject = ESCOLA]	Variance	,000	,000
repro_dummie [subject = ESCOLA]	Variance	6,141	4,691	1,309	,190	1,374	27,443
trabalha_fora [subject = ESCOLA]	Variance	12,293	9,054	1,358	,175	2,902	52,071

Fonte dos dados das saídas: INEP/SAEB (2018). Elaboração do modelo: própria.