

Determinantes do Desempenho no ENEM na Região Nordeste: Uma Análise de Dados em Painel do Período de 2015 a 2019

Adeilson Elias de Souza¹  | Luiza Mikaela de Sá Santos²  | Igor de Menezes Larruscaim³  | Cássio da Nobrega Besarria⁴ 

¹ Universidade Federal da Paraíba. E-mail: adeilsonelias@outlook.com

² Universidade Federal da Paraíba. E-mail: luizamikaela@hotmail.com

³ Universidade Federal do Rio Grande do Sul. E-mail: igor_menezes06@hotmail.com

⁴ Universidade Federal da Paraíba. E-mail: cassiodanobrega@yahoo.com.br

RESUMO

O objetivo do trabalho é avaliar como o desempenho dos candidatos do ENEM, na Região Nordeste do Brasil, é influenciado pelos diversos fatores ligados às características sociais e econômicas em que estão inseridos no período de 2015 a 2019. Para isso, foram coletadas informações do INEP referentes ao ENEM que possibilitaram estimar modelos de dados em painel. Os resultados foram condizentes com a literatura econômica e mostraram que principalmente a escolaridade dos pais, o tipo de escola e a disposição de ferramentas, como o celular e o computador, e acesso à internet são fatores determinantes para os candidatos situados na Região Nordeste aumentarem sua média do ENEM. Por outro lado, fatores como a idade do candidato e se ele pertence à classe econômica de baixa renda podem reduzir suas médias no ENEM.

PALAVRAS-CHAVE

Desempenho no ENEM, Nordeste, Dados em painel

Determinants of ENEM Performance in the Northeast Region: a Panel Data Analysis from 2015 to 2019

ABSTRACT

The objective of this work was to evaluate how the performance of ENEM candidates, in the Northeast Region of Brazil, is influenced by the various factors linked to the social and economic characteristics in which they are inserted in the period from 2015 to 2019. For this, information was collected from INEP regarding ENEM, which made it possible to estimate panel data models. The results were consistent with the economic literature and showed that, mainly, the parents' education, the type of school and the availability of tools, such as cell phones and computers, and internet access are determining factors for candidates located in the Northeast Region to increase their exam average. On the other hand, factors such as the candidate's age and whether he belongs to the low-income economic class can reduce his averages in the ENEM.

KEYWORDS

ENEM performance, Northeast, Panel data

CLASSIFICAÇÃO JEL

I2, I21, I24, C3, C33

1. Introdução

O tema educação é de importância fundamental para a compreensão de diversos fenômenos econômicos como o crescimento da renda, desigualdade e desenvolvimento humano de uma sociedade. Como destaca Cunha et al. (2006), a educação é um dos pilares da economia, porque possibilita o aprimoramento das capacidades técnicas dos indivíduos e, conseqüentemente, contribui para o progresso tecnológico. Além desses benefícios, ao possibilitar o desenvolvimento do pensamento crítico e reflexivo, gera um papel civilizatório na sociedade.

No Brasil, uma das principais ferramentas de avaliação da qualidade de ensino é o Exame Nacional do Ensino Médio (ENEM). A prova tem como objetivo identificar o nível de aprendizado dos participantes, a fim de torná-los aptos a frequentar cursos de nível superior nas universidades brasileiras, de acordo com o seu desempenho. Em 2019, mais de 5,1 milhões de inscrições foram realizadas, com uma taxa de participação de 72,81%. Em relação aos resultados, observou-se que a média das notas caiu nas quatro áreas do conhecimento em relação a 2018, com reduções respectivas de 2,37% em Matemática; 12,04%, em Ciências Humanas; 1,15%; em Linguagens e Códigos; e 3,34%, em Ciências da Natureza (INEP, 2019).

Os prestadores do exame podem utilizar a nota do ENEM para ter acesso ao ensino superior público, por meio do Programa Sistema de Seleção Unificada (SISU), bem como ao ensino superior privado, sobretudo o Fundo de Financiamento ao Estudante do Ensino Superior (FIES), que funciona como um intermediário de financiamento estudantil para os alunos que desejam ingressar em faculdades particulares, e o Programa Universidade para Todos (Prouni), que distribui bolsas estudantis para alunos que não podem pagar uma universidade particular, que vão de 50% a 100%, para alunos que não possuem condições de pagar uma universidade particular. No entanto, um dos requisitos para os alunos participarem desses processos, no que tange ao desempenho, é de que a média de suas notas seja igual ou superior a 450 pontos e tenham nota maior que zero na redação (INEP, 2019).

Visto que o Brasil apresenta desigualdades em diversos aspectos socioeconômicos, os resultados de exames como o ENEM são uma forma de avaliar as disparidades presentes no sistema educacional. De acordo com Netto e Figueiredo (2010), existe uma disparidade regional considerável na qualidade e acesso à educação. Como no caso da Região Nordeste, que apresentou o maior índice de desigualdade educacional no Brasil durante o período de 1986 a 2005. Quanto aos resultados do ENEM, essa heterogeneidade na qualidade do ensino também é identificada na região. Lobo et al. (2017) identificaram que no ano de 2013, na Região Nordeste, o desempenho médio na prova de matemática do ENEM era afetado positivamente por fatores como acesso à boa estrutura escolar, o candidato ter maior renda, os pais terem maior escolaridade, acesso à internet e o fato de serem brancos e do sexo masculino. Em contrapartida, o fato de o candidato ter baixa renda, ter trabalhado na adolescência e ter estudado

em escola pública influenciavam negativamente o seu desempenho médio.

Considerando esse contexto, este estudo teve como objetivo geral avaliar como o desempenho dos candidatos do ENEM, na Região Nordeste do Brasil, é influenciado pelos diversos fatores ligados às características sociais e econômicas em que estão inseridos no período de 2015 a 2019. Para isso, elencamos os seguintes objetivos específicos: caracterizar o perfil dos estudantes que se submetem ao teste do ENEM na Região Nordeste; identificar os fatores que influenciam, de forma mais expressiva, o desempenho no ENEM nessa região e levantar dados que possibilitem apontar indicadores a serem trabalhados em políticas educacionais com vistas a melhorar a qualidade do sistema de ensino e, conseqüentemente, aprimorar o aproveitamento no referido exame. O estudo limita sua análise até o ano de 2019 para que não sejam captados os possíveis efeitos da pandemia da COVID-19 nos resultados. Uma análise mais aprofundada que trate detalhadamente dessa questão é tema para um estudo futuro, já que alguns dos efeitos da pandemia ainda podem ser sentidas na estrutura socioeconômica do país.

Os aspectos metodológicos visam adaptar a proposta de Figueiredo et al. (2014), e buscam entender os fatores que afetam o desempenho dos candidatos que prestaram o ENEM em um período mais recente. Para que fosse possível captar as diferentes características dos candidatos ao longo do período de tempo determinado se utilizou o método de dados em painel. A adoção dessa estratégia empírica se diferencia da maioria dos trabalhos sobre o tema desempenho educacional, visto que geralmente essas análises são do tipo cross-section e avaliam apenas um ano específico.

A escolha desse tema e do recorte regional se justifica pela importância do tema educação para o entendimento do crescimento e do desenvolvimento econômico e humano, assim como pelo contexto educacional do Nordeste, que apresenta indicadores preocupantes, como o maior índice de desigualdade educacional entre as regiões do país, o que se reflete nos indicadores do ENEM. Além disso, não existem uma grande variedade de estudos mais detalhados sobre a qualidade da educação nessa região. Portanto, compreender bem mais os determinantes do desempenho escolar no Nordeste é de suma importância para a orientação de políticas educacionais que visem melhorar a qualidade educacional e repercutam no resultado escolar do aluno. Logo, espera-se contribuir para ampliar a literatura a respeito do tema, uma vez que o estudo possui a particularidade de ser exclusivo para a referida região sob o período de análise e a estratégia empírica adotada.

Este estudo está estruturado em mais quatro seções, além desta introdução. Na segunda, apresenta-se a revisão da literatura; na terceira, a metodologia empregada nas estimações do modelo da pesquisa; na quarta, os resultados obtidos; e na quinta, as considerações finais.

2. Revisão da literatura

Na literatura econômica acerca do desempenho escolar, a qualidade da educação é mensurada considerando-se diferentes fatores, como: frequência, atraso, evasão escolar e proficiência em testes, como é o caso do ENEM no Brasil. Conhecer os condicionantes do desempenho escolar é de suma relevância para direcionar políticas educacionais voltadas para melhorar a qualidade do sistema de ensino.

Em âmbito internacional e nacional, não é difícil encontrar evidências empíricas exclusivamente sobre os fatores que explicam o desempenho escolar do aluno. Conforme Soares (2004), tais fatores podem ser reunidos em três principais grupos: características individuais dos alunos, da família e da escola.

A literatura internacional inclui o estudo de Levacic et al. (2005), que analisou os efeitos dos recursos escolares no desempenho educacional na Inglaterra, utilizando os métodos de Mínimos Quadrados Ordinários (MQO) e de Variáveis Instrumentais (VI). Os resultados encontrados evidenciaram que aumentos nas despesas por aluno bem como a redução da proporção de alunos por professor afetaram positivamente os resultados em Matemática e em Ciências de alunos com idades entre 14 e 15 anos, em 2003. Os autores também concluíram que, quando esses efeitos são estimados por VI são, cerca de, até dez vezes maiores do que os encontrados por MQO, o que indica que essa relação pode ser subestimada quando a endogeneidade não é levada em consideração. Esses recursos escolares exercem mais efeitos nos alunos de famílias mais pobres.

Fazendo essa mesma análise, contudo, por meio de equações simultâneas multinível, Steele et al. (2007) encontraram resultados parecidos e apontaram uma associação positiva entre os recursos escolares e o desempenho nas disciplinas 'Matemática' e 'Ciências' e estimaram que um aumento de £ 1000 com despesas por aluno aumentaria o desempenho escolar em 0,07, 0,2 pontos, aproximadamente, em Matemática e Ciências, respectivamente, e que uma redução na proporção de alunos por professor levaria a um ganho no rendimento acadêmico de 0,12 e 0,13 pontos para essas disciplinas.

Quanto à literatura nacional, entre os trabalhos que focam a importância das características individuais para explicar o desempenho, estão os estudos que apontam que o principal determinante é a própria habilidade cognitiva do aluno. Nessa linha, Queiroz et al. (2018) e Di Pietra et al. (2018) constataram que estudantes com mais habilidades cognitivas encontram menos barreiras no processo de ensino/aprendizagem. Isso implica atingir mais anos de estudo, o que reflete nos resultados socioeconômicos individuais e no desenvolvimento econômico do país. Por outro lado, Hanushek (1979) aborda que é difícil obter uma medida adequada para a variável habilidade inata/capacidade de aprendizagem.

Outra linha de argumento assume que a quantidade de horas de estudo fora da

escola é um fator determinante no desempenho do aluno, porque o fato de ter que se dedicar ao estudo fora do ambiente escolar influencia positivamente seu desempenho nas avaliações (Jesus e Laros, 2004; Macedo, 2004). Os autores também notaram que a trajetória escolar influencia o desempenho dos alunos e que fatores presentes especialmente nas escolas públicas do país, como a repetência, a evasão e o atraso escolar, participam diretamente do rendimento do aluno.

Fatores individuais como cor/raça e sexo também podem estar associados ao desempenho do aluno. De acordo com Andrade (2012), a cor do aluno, geralmente, demonstra sua herança cultural, o que pode explicar seu desempenho. A variável 'cor do aluno' também é influenciada significativamente pela variável 'renda', porquanto reflete um cenário em que os salários das pessoas não brancas são menores do que os das brancas. Isso explica o desempenho superior dos alunos brancos em relação aos pardos e aos pretos.

Montalvão e Neubert (2016) verificaram que estudantes negros apresentam trajetórias escolares mais irregulares do que os brancos. E mesmo quando indivíduos negros dispõem de recursos econômicos que os coloquem em vantagem relativa, o alcance e a trajetória educacional ainda são piores do que os dos indivíduos brancos. Isso sugere a importância da raça/cor como mecanismo gerador de desigualdades educacionais no Brasil. Existe uma ampla linha de pesquisa que aponta que a determinação do desempenho escolar é influenciada, essencialmente, por fatores associados às características familiares, como: renda familiar, escolaridade dos pais, número de filhos na família; gênero do chefe de família e estrutura familiar.

Menezes e Soares (2010), ao examinar os possíveis condicionantes dos resultados em Matemática de estudantes do ensino fundamental e do médio, usando microdados do SAEB de 2003, encontrou que, além dos atributos individuais dos alunos, as variáveis mais importantes para explicar os resultados escolares são referentes aos atributos da família, com destaque para a escolaridade da mãe, a quantidade de livros e o trabalho fora de casa, sendo que esta última tem efeito negativo no desempenho em Matemática. Os resultados também apontaram para diferenças de gêneros, já que os homens, em média, têm melhores desempenhos em Matemática quando comparados com as mulheres em todos os níveis de escolaridade.

Melo e Arakawa (2012) analisaram as desigualdades regionais das influências do *background* familiar sobre o desempenho escolar no ENEM de 2008 de estudantes com idades entre 15 e 20 anos que estavam concluindo o ensino. Para isso, utilizaram o método de MQO. Os resultados mostraram que existe uma relação positiva entre a educação e a renda dos pais sobre os resultados potenciais dos filhos, em todas as regiões do Brasil. Também constaram que essa relação entre o *background* familiar e o desempenho no ENEM é diferente nas regiões do país, pois é mais forte nas regiões menos desenvolvidas do que nas mais desenvolvidas.

Mendes e Karruz (2012) também investigaram os possíveis efeitos do *background*

familiar nos resultados de alunos no ENEM 2012 nas grandes regiões do Brasil. Estimando uma regressão múltipla por MQO, os autores constataram que quanto maior a escolaridade dos pais e a faixa de renda familiar, melhor será o desempenho dos estudantes no ENEM. Além disso, encontraram diferenças inter-regionais das influências da educação e da renda dos pais sobre o desempenho escolar. As regiões Centro-Oeste e Sul registraram, respectivamente, maior influência da educação da mãe e do pai. Apontaram também que a renda familiar exerce mais e menos influência no Nordeste e no Sul, respectivamente.

A literatura referente aos atributos da escola assegura que o desempenho dos alunos nas avaliações também se deve à infraestrutura da escola, às características dos docentes, à qualidade da educação, à dependência administrativa e à localização da escola, entre outros. Nesse sentido, destaca-se o trabalho de Albernaz et al. (2002), que, ao estimar modelos hierárquicos lineares com base em microdados do SAEB de 1999 para a oitava série, concluíram que aproximadamente 80% da variância de desempenho médio entre as escolas se devem a diferenças na composição socioeconômica de seus alunos. A rede a que a escola pertence está associada de forma importante ao desempenho de seus alunos, em que o desempenho médio da escola particular supera o da escola pública. Outras variáveis escolares, como a escolaridade dos professores e a qualidade da infraestrutura física da escola, também influenciam o rendimento médio dos estudantes de forma significativa.

Contribuindo com a hipótese do efeito da dependência administrativa, Menezes-Filho (2007) mostrou que os alunos das escolas privadas apresentam um desempenho melhor do que os alunos das escolas públicas, embora controlado o efeito de variáveis de *background* familiar. Como visto, há uma diversidade de fatores que podem estar associados aos resultados dos alunos em avaliações de larga escala. Nesse sentido, alguns trabalhos testaram um conjunto dos fatores apontados como determinantes do desempenho educacional.

Figueiredo et al. (2014) avaliaram como as circunstâncias sociais influenciam o desempenho escolar dos alunos que prestam o exame ENEM. Para isso, utilizaram, inicialmente, uma abordagem não paramétrica assumindo o *Roemer Identification Axiom*, isto é, a independência entre as circunstâncias e o esforço. Em seguida, relaxaram essa hipótese buscando mensurar o viés da omissão do esforço e do talento individual. O primeiro conjunto de resultados indicou grande diferença no esforço empregado entre os tipos para obter um bom desempenho no ENEM. O segundo concluiu que, quanto melhor o conjunto de oportunidades expresso no *background* familiar, melhores são os desempenhos no ENEM. Também se notou um efeito positivo nos indivíduos que moram em área urbana, são do sexo feminino, não negros, que estudaram em escolas privadas e de boa qualidade.

Lobo et al. (2017) analisaram o desempenho dos alunos que estavam cursando o terceiro ano do ensino médio na prova de matemática do ENEM, para a Região Nordeste, no ano de 2013. Os autores utilizaram um modelo de regressão linear

hierárquico de dois níveis. O primeiro nível do modelo corresponde ao efeito das várias relativas ao *background* familiar do aluno. Enquanto o segundo nível representa o tipo de escola em que o aluno estudou. A partir dessa análise os autores identificaram que a diferença entre o tipo de escola, pública ou privada, é o principal fator de explicação da variância total das notas, representando 15,28%. Verificou-se que para a escola pública houve uma redução aproximada de 55,5 pontos na média do estudante. Além disso, fatores relacionados ao *background* familiar do aluno, como ter trabalhado na adolescência, ser de baixa renda e ter um número maior de pessoas morando na mesma casa, contribuem para a redução da nota média deste.

Por sua vez, Araújo (2019) estudou os determinantes do desempenho no ENEM de alunos com idades entre 15 e 25 anos da cidade mineira de Viçosa, entre os anos de 2015 e 2017. O autor utilizou um modelo de regressão hierárquica em três níveis com intercepto aleatório para captar a correlação entre estudantes da mesma escola. Esses níveis são referentes ao aluno, à escola e ao ano. Como principais resultados, o autor verificou que, no modelo nulo, a variação entre as escolas explica 60% do desempenho educacional dos alunos. Quando inseridas as variáveis de controle, tal índice se reduz para 5,4% e cai o percentual da variância total explicada pelo tipo de escola, mostrando a importância do nível do aluno para explicar o desempenho escolar. Entre os fatores que explicam o desempenho do aluno, estão a idade, a raça e o sexo. Em relação aos fatores que medem o *background* familiar, destacam-se a renda da família, a alta escolaridade da mãe e o acesso à internet. Ressalta-se que a dependência administrativa das escolas é o principal condicionante para o desempenho dos alunos.

Portanto, observa-se que dentre os fatores determinantes para o desempenho no exame do ENEM, destacam-se: i) as desigualdades raciais, que afetam de forma negativa o rendimento dos candidatos não brancos, como apontado pelos trabalhos de Andrade (2012); Montalvão e Neubert (2016) e Figueiredo et al. (2014); ii) a renda familiar como propulsora do rendimento, segundo Melo e Arakawa (2012) e Mendes e Karruz (2012) e; iii) o candidato pertencer a escola privada contribui de forma positiva para uma maior pontuação no ENEM, como salientado Menezes-Filho (2007); Lobo et al. (2017) e Araújo (2019).

3. Metodologia

Esta pesquisa é caracterizada como quantitativa, porquanto propôs estimar os efeitos que os fatores sociais e econômicos dos candidatos exercem no desempenho do ENEM, e exploratória e descritiva, visto que analisou as características dos participantes no exame referido.

As subseções 3.1 e 3.2 abordam, respectivamente, o procedimento estatístico e econométrico, juntamente com a especificação do modelo disposto na pesquisa, e a natureza dos dados.

3.1 Procedimento Estatístico e Econométrico

Em relação aos procedimentos⁵ metodológicos aplicados no trabalho, calculou-se a estatística descritiva da proporção de participantes, pontuação média no exame, desvio padrão, menor e maior desempenho no ENEM para variáveis sociais e econômicas referentes aos participantes. i) as variáveis sociais englobaram: sexo, tipo de ensino, zona de residência, cor e se o pai e/ou a mãe concluiu o ensino médio; ii) enquanto as variáveis econômicas foram: ser caracterizado como baixa renda, possuir celular, ter acesso a internet e possuir computador.

Tendo realizado os procedimentos estatísticos, estima-se o modelo do trabalho por meio de dados em painel para o período de 2015 a 2019, visando a obtenção de respostas que permeiem o rendimento dos participantes do ENEM a partir de fatores sociais e econômicos elencados na pesquisa.

De acordo com Greene (2007), por meio do método de dados em painel, os cientistas podem identificar mudanças de padrões resultantes dos agentes da sociedade ao longo do tempo. Marques (2000) e Baltagi (2005) asseveram que, usando o método de dados em painel, podem-se estimar e identificar fatores não presentes em cross-section ou série temporal e um número elevado de graus de liberdade acompanhado de mais eficiência dos estimadores.

Utilizando uma notação clássica do modelo de regressão linear, que considera os dados apenas em corte transversal, temos um modelo representativo como:

$$Y_i = \alpha + \beta X_i + \epsilon_i \quad (1)$$

Em que Y_i seria a média do ENEM em um ano representativo, α corresponde ao intercepto, β são os coeficientes das variáveis explicativas em relação à média do ENEM, X_i o conjunto de variáveis sociais e econômicas que afetam o rendimento no ENEM e ϵ_i o termo de erro do modelo.

Com a introdução do tempo na equação (3.1), obtêm-se uma representação de um modelo de dados em painel, definido na equação (3.2) como:

$$Y_{it} = \alpha + \beta X_{it} + \epsilon_{it} \quad (2)$$

Em que os subscritos i e t estão associados as observações distribuídas entre os indivíduos no tempo, neste caso, os indivíduos que realizaram o exame ENEM no período de 2015 a 2019.

No método de dados em painel, existem algumas especificações que determinam a identificação do seu formato. Segundo Greene (2007), são elas: efeitos aleatórios, efei-

⁵Todos os procedimentos estatísticos e econométricos foram realizados via software R.

tos fixos e regressão *Pooled*. Neste último, a estimação em MQO produz estimadores consistentes e eficientes que possibilitam uma relação constante ao longo do tempo por parte dos indivíduos. Porém, caso exista fatores não observados, dispostos em ϵ_{it} que não seja aleatório, mas sistemático, a regressão *pooled* se torna inapropriada.

Outra especificação é a dos estimadores considerando efeitos fixos, que Segundo Wooldridge (2010, p. 450), “levam em conta uma correlação arbitrária entre a_1 e as variáveis explicativas em qualquer período de tempo, como em primeira diferença. Por esse motivo, qualquer variável explicativa que seja constante, ao longo do tempo, para todo i é removida pela transformação de efeitos fixos”. Nesse sentido, uma variável como a distância em quilômetros entre a residência e a escola não deve ser inclusa nesse caso.

Por outro lado, a inviabilidade de testar variáveis como a distância em quilômetros entre a residência e a escola, que não mudam com o tempo, pode ser contornada por meio da regressão por efeitos aleatórios, ao permitir a inclusão de variáveis que não se alteram ao decorrer do tempo.

O processo de escolha entre qual destes efeitos adotar na pesquisa deve levar em conta o teste F de Chow, que compara se existe diferenças significativas entre o modelo *Pooled* e efeitos fixos. Ademais, têm-se o teste de Breusch-Pagan LM, que defronta o *Pooled* e efeitos aleatórios, sendo na hipótese nula a indicação que a variância por indivíduo é zero, recomendando o *Pooled*.

Podem ser observadas algumas diferenças entre os modelos de efeitos fixos e aleatórios. Enquanto o primeiro avalia as distinções nos interceptos, devido à inclinação e aos elementos que constam na amostra, o segundo não apresenta correlação entre os regressores e os efeitos individuais e a variância do termo de erro obtida para cada grupo de elemento da amostra. O teste de Hausmann é considerado determinante para explicar o tipo de efeito que será adotado na pesquisa. Na equação (3.3), é possível observar a formulação matemática do teste obtido no trabalho de Hausman (1978).

$$\xi_H = (\hat{\beta}_{EF} - \hat{\beta}_{EA})' [\hat{V}(\hat{\beta}_{EF}) - \hat{V}(\hat{\beta}_{EA})]^{(-1)} (\hat{\beta}_{EF} - \hat{\beta}_{EA}) \quad X^{(K)} \quad (3)$$

Aqui, $\hat{\beta}_{EF}$ representa os coeficientes gerados considerando efeitos fixos, $\hat{\beta}_{EA}$, considerando efeitos aleatórios, tendo em vista a exclusão da constante, $(\hat{\beta}_{EF} - \hat{\beta}_{EA})'$ representa o transposto do vetor de diferença dos modelos dispostos, $[\hat{V}(\hat{\beta}_{EF}) - \hat{V}(\hat{\beta}_{EA})]^{(-1)}$ a inversa da matriz de diferença das variâncias dos modelos, \hat{V} , o coeficiente de estimação das matrizes de variância e covariância, e ξ_H denota a distribuição qui-quadrado.

Segundo Greene (2007, p. 208), o teste de Hausman tem na hipótese nula o entendimento de que o modelo de efeitos aleatórios é mais adequado que o de efeitos fixos. Nesse sentido, a hipótese é de que, como os efeitos individuais são correlacio-

nados com os regressores, a estimação por efeitos aleatórios se torna problemática, e o modelo com efeitos fixos passa a ser o mais recomendado.

Na Tabela 1 são expostas as hipóteses dos testes F de Chow, que compara e indica qual o modelo preferível entre o de efeitos fixos e o *Pooled*, o de Breusch-Pagan LM, para efeitos aleatórios ou *Pooled*, e o de Hausman, para efeitos fixos ou aleatórios.

Tabela 1. Testes de agrupamento dos dados

Testes	Hipóteses	
	H0	H1
F de Chow	<i>Pooled</i>	Efeitos fixos
Breusch-Pagan LM	<i>Pooled</i>	Efeitos aleatórios
Hausman	Efeitos aleatórios	Efeitos fixos

Fonte: Elaboração dos autores.

Após a escolha dos efeitos e estimação do modelo, são realizados os testes do multiplicador de Lagrange (LM), de Breusch-Godfrey/Wooldridge, de Jarque-Bera e de Breusch-Pagan aplicados nos resíduos dos modelos estimados. Na Tabela 2 pode-se acompanhar as hipóteses nulas e alternativas destes testes.

Tabela 2. Testes utilizados nos resíduos do modelo e suas respectivas hipóteses

Testes	Hipóteses	
	H0	H1
LM	$\beta_1 \dots \beta_n = 0$	pelo menos algum β será diferente de 0
Breusch - Godfrey/Wooldridge	ausência de autocorrelação serial	presença de autocorrelação serial
Jarque - Bera	os erros se distribuem normalmente	a distribuição dos erros não é normal
Breusch - Pagan	homocedasticidade	heterocedasticidade

Fonte: Elaboração dos autores.

O teste do multiplicador de Lagrange possui como hipótese nula a indicação de que todos os parâmetros β são iguais a zero em detrimento de que pelos menos algum dos estimadores do modelo seja estatisticamente diferente de zero como hipótese alternativa. O teste de autocorrelação serial de Breusch-Godfrey/Wooldridge verifica se os resíduos possuem ausência de autocorrelação serial como hipótese nula e presença de autocorrelação serial, ou seja, os erros são autocorrelacionados no tempo como hipótese alternativa. O teste de Jarque Bera observa se os resíduos possuem distribuição normal como hipótese nula e ausência de distribuição normal como hipótese alternativa. E por fim, no teste de Breusch-Pagan observamos se os erros são homocedásticos como hipótese nula e se são heterocedásticos como hipótese alternativa.

Se tratando do modelo econométrico estimado, a pesquisa visou analisar como os diversos fatores sociais e econômicos podem afetar o rendimento dos participantes do Exame Nacional do Ensino Médio (ENEM). Seu modelo é uma adaptação do que foi proposto no trabalho de Figueiredo et al. (2014), que aborda como as diversas condições apresentadas pelos estudantes afetam seus desempenhos no ENEM, apresentado na equação 3.4:

$$M_{notaenem} = \beta_0 + \beta_1 idade + \beta_2 idade^2 + \beta_3 sexo + \beta_4 raça + \beta_5 tipo de escola + \beta_6 ESC mãe + \beta_7 ESC pai + \beta_8 moradores na casa + \beta_9 localização + \beta_{10} baixa renda + \beta_{11} celular + \beta_{12} internet + \beta_{13} computador + \varepsilon \quad (4)$$

Nessa equação, $M_{NotaEnem}$ representa a variável dependente do modelo, definida como a média das notas obtidas nas áreas de Ciências da Natureza, Ciências Humanas, Linguagens e Códigos, Matemática e Redação para cada candidato; β_0 representa a constante do modelo; $\beta_1, \dots, \beta_{13}$ são os parâmetros estimados; quanto às variáveis explicativas, tem-se: idade, que se refere à idade do participante; $idade^2$, que é a idade do participante ao quadrado; sexo, ao gênero masculino ou feminino; raça indica se o aluno é branco ou o contrário; o tipo da escola, se é pública ou privada; Esc mãe e Esc pai, a escolaridade da mãe e do pai, respectivamente; baixa renda, se o aluno percebe até três salários mínimos; moradores na casa diz respeito à quantidade de pessoas que residem no domicílio; celular, para identificar se possui este aparelho; internet, para participantes que dispõem desse serviço; computador, para pessoas que possuem esse equipamento; ε_i é o termo de erro do modelo.

A partir do Quadro 1 podemos acompanhar as variáveis e os seus respectivos sinais esperados.

Quadro 1. Variáveis e sinais esperados

Variáveis	Sinais	Autores
Idade	(-)	Silva (2017)
Sexo	(-)	Torres <i>et al.</i> (2020)
Raça	(+)	Figueirêdo <i>et al.</i> (2014), Torres <i>et al.</i> (2020) e Silva (2017)
Tipo da escola	(+)	Figueirêdo <i>et al.</i> (2014) e Torres <i>et al.</i> (2020)
Escolaridade do pai	(+)	Figueirêdo <i>et al.</i> (2014) e Torres <i>et al.</i> (2020)
Escolaridade da mãe	(+)	Figueirêdo <i>et al.</i> (2014) e Torres <i>et al.</i> (2020)
Baixa renda familiar	(-)	Figueirêdo <i>et al.</i> (2014), Torres <i>et al.</i> (2020) e Silva (2017)
Localização	(+)	Figueirêdo <i>et al.</i> (2014)
Moradores na casa	(-)	Lobo <i>et al.</i> (2017)
Celular		
Internet	(+)	Lobo <i>et al.</i> (2017), Torres <i>et al.</i> (2020) e Silva (2017)
Computador	(+)	Lobo <i>et al.</i> (2017)

Fonte: Elaboração dos autores.

A partir do Quadro 1, nota-se que de acordo com Silva (2017) a idade é um fator que contribui negativamente com a média no ENEM, ao representar alunos que possivelmente apresentaram evasão escolar. Segundo Torres *et al.* (2020), os indivíduos do sexo masculino apresentam rendimento superior ao feminino. De acordo com Figueiredo *et al.* (2014), Torres *et al.* (2020) e Silva (2017), os candidatos de cor branca apresentam desempenho superior aos não brancos, assim como os de escolas particulares em relação aos de escola pública e os localizados na zona urbana em relação a rural, por outro lado, os candidatos de baixa renda apresentam desempenho

inferior aos candidatos com renda superior. Segundo Figueiredo et al. (2014) e Torres et al. (2020), o pai e a mãe possuem ensino médio apresentam uma relação positiva com o rendimento no exame. E por fim, de acordo com Lobo et al. (2017), o número de moradores na casa reduz o desempenho no ENEM, enquanto possuir internet e computador em casa contribui com a elevação na pontuação média no exame.

3.2 Natureza dos dados

Os dados dispostos na pesquisa foram coletados no site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) referentes ao ENEM, no período de 2015 a 2019. A base de dados citada aborda informações sociais e econômicas dos candidatos inscritos nesse exame. Convém enfatizar que só foram considerados os candidatos da Região Nordeste que compareceram à prova e que não zeraram a média das quatro áreas do conhecimento e da redação. Estes últimos foram excluídos por causa do elevado índice de abstenção, que apresentou uma média de, aproximadamente, 29,99%, no período de 2015 a 2019, de acordo com dados do INEP (2019).

Na estratégia empregada para construir a amostra, a escolha pela Região Nordeste teve em vista o histórico de relativo atraso socioeconômico. Apesar de evoluções em alguns indicadores, cabe analisar possíveis alternativas que visem o desenvolvimento da região. De acordo com dados do IBGE (2019), o Nordeste em 2019, com o valor de 0,543, possuiu o pior índice de Gini dentre todas as regiões nacionais, indicando uma elevada concentração de renda.

Além disso, de acordo com Netto e Figueiredo (2010) a região possuía o maior índice de desigualdade educacional no Brasil, considerando o período de 1986 a 2005. Segundo Rodrigues et al. (2017), analisando os dados do Censo Demográfico de 2010, aproximadamente 60% da população adulta dos estados do Nordeste sequer havia concluído o ensino fundamental. Até esse período os estados da região continuavam apresentando alguns dos piores índices de desigualdade educacional do país.

A amostragem do presente estudo corresponde aos anos 2015 a 2019, quando foi reformulada a isenção da taxa de inscrição, a qual passou a valer apenas para candidatos que concluíram o ensino médio em escola pública e que comprovem carência², até o ano que antecede a pandemia, buscando isolar possível viés.

O Quadro 2 retrata as variáveis da pesquisa e as suas respectivas descrições.

Realizados os filtros e excluídos os *missing*, a amostra ficou composta de 920.127 observações, sendo 211.363 no ano de 2015; 224.919, em 2016; 176.434, em 2017; 155.081, em 2018; e 152.330, em 2019. Portanto, houve uma queda no número de alunos que realizaram a prova do ENEM ao longo dos últimos anos. Destaca-se que as variáveis sexo, raça, tipo da escola, escolaridades do pai e da mãe, baixa renda, localização, moradores na casa, celular, internet e computador são expressas como

²Esse fato teve o objetivo de reduzir o número de inscrições de candidatos desinteressados com o exame proposto.

Quadro 2. Descrição das variáveis da pesquisa

Variáveis	Descrição
Rendimento ENEM	Corresponde às médias individuais das notas obtidas pelos participantes do exame ENEM
Idade	Representa a idade do participante em valor numérico
Sexo	1 - Feminino 0 - Masculino
Raça	1 - Branco 0 - Não branco
Tipo da escola	1 - Privada 0 - Pública
Escolaridade do pai	1 - Caso o pai tenha concluído ao menos o ensino médio 0 - Caso contrário
Escolaridade da mãe	1 - Caso a mãe tenha concluído ao menos o ensino médio 0 - Caso contrário
Baixa renda	1 - 1 - Caso a renda da família seja até três salários mínimos 0 - Caso contrário
Localização	1 - Urbana 0 - Rural
Moradores na casa	Retrata a quantidade de pessoas que residem na casa onde o participante reside
Celular	1 - Possui celular 0 - Caso contrário
Internet	1 - Possui acesso à internet na residência 0 - Caso contrário
Computador	1 - Possui computador na residência 0 - Caso contrário

Fonte: Elaboração dos autores a partir dos dados da pesquisa.

dummies para a realização das estimações do modelo.

4. Resultados

Nas subseções 4.1, 4.2 e 4.3, são apresentados e discutidos, respectivamente, a estatística descritiva das variáveis explicativas *dummies*, consideradas neste trabalho, em relação à nota média do ENEM; os resultados dos testes; e os resultados da estimação do modelo de dados em painel para verificar a relação existente entre as variáveis explicativas e a dependente.

4.1 Estatística descritiva

Partindo da proposta de caracterizar o perfil dos estudantes que se submeteram ao exame do ENEM, no período de 2015 a 2019, na Região Nordeste, e de identificar os diferenciais da nota média dos diferentes participantes do ENEM, a Tabela 3 apresenta o percentual de participantes do ENEM e a pontuação média de acordo com as variáveis consideradas neste estudo. A Tabela 4 apresenta a pontuação média conforme a soma de algumas das características socioeconômicas representadas nas variáveis, por exemplo: homem branco que não tem baixa renda e estudou em escola privada; mulher não branca com baixa renda e estudou em escola pública, etc.

Conforme as informações apresentadas na Tabela 3, as candidatas do sexo feminino apresentam uma participação consideravelmente maior na amostra em relação aos alunos do sexo masculino, com 58,7% e 41,3%, respectivamente. Entre esses

Tabela 3. Estatísticas descritivas: Pontuação média de acordo com a variável

Variáveis	% de participantes	Média	Desvio	Mínimo	Máximo
Feminino	58,7%	497,0	75,3	48,0	834,0
Masculino	41,3%	508,0	82,0	40,0	880,0
Privado	16,4%	583,0	85,1	140,0	834,0
Público	83,6%	481,0	71,6	86,7	824,0
Branco	22,3%	521,0	85,4	83,2	836,0
Não Branco	77,7%	496,0	74,9	40,0	880,0
Pai concluiu ao menos o EM	33,8%	536,0	84,7	81,7	836
Caso contrário	66,18%	487,0	70,1	48,0	880
Mãe concluiu ao menos o EM	44,7%	528,0	83,0	48	880
Caso contrário	55,3%	483,0	68,2	68	824
Renda baixa	88,0%	491,0	71,0	40,0	880
Caso contrário	12,0%	579,0	85,3	81,7	836
Urbana	95,9%	499,0	83,4	86,7	834,0
Rural	4,1%	472,0	71,2	172,0	733,0
Possui celular	97,3%	502,0	78,3	40,0	880,0
Não possui	2,7%	470,0	72,6	116,0	785,0
Possui Internet	61,6%	518,0	81,1	60,0	880,0
Não possui	38,4%	475,0	65,4	40,0	830,0
Possui computador	61,6%	526,0	81,5	48,0	880,0
Não possui	38,4%	477,0	66,5	40,0	830,0

Fonte: Elaboração dos autores a partir dos dados da pesquisa.

grupos, observa-se ligeira diferença na pontuação média, equivalente a 11 pontos a mais para os do sexo masculino. Quanto à raça, os não brancos representam 77,7% dos que se submeteram ao Exame, e 22,3% brancos. Em termos de desempenho, os brancos obtiveram uma pontuação média maior, correspondente a 25 pontos a mais.

No que se refere ao tipo de escola, 83,6 % dos participantes estudaram em escola pública, enquanto apenas 16,4% estudaram em escola privada. Por outro lado, os estudantes de escolas particulares obtiveram melhor desempenho médio (583,0) em relação aos de escolas públicas (481,0). Sobre a escolaridade do pai e da mãe, apenas 33,8% e 44,7%, respectivamente, dos pais e das mães dos participantes concluíram, ao menos, o Ensino Médio. Ademais, as maiores pontuações médias foram para esses grupos, se comparados com os dos pais e das mães que não concluíram sequer o Ensino Médio.

Quanto à renda baixa, é notória uma concentração de 88% da amostra no grupo de participantes que têm até três salários mínimos, isto é, cuja renda é baixa. Para a pontuação média, nota-se que os que não têm renda baixa obtiveram 579 pontos, equivalente a 88 pontos a mais do que os de renda baixa.

No tocante à localização da escola, quase toda a amostra (95,9%) é composta de participantes que residem em área urbana. Eles apresentaram uma pontuação mé-

dia de 499,0, enquanto para os residentes em área rural é de 422,0. Em relação às variáveis celular, internet e computador, existe uma representação maior na amostra de participantes que possuem esses equipamentos, os quais apresentam maior pontuação média. A Tabela 4 apresenta a média das notas de acordo com algumas das características socioeconômicas do candidato.

Tabela 4. Pontuação média de acordo com as características socioeconômicas

Variável		Pontuação média			
		Homem branco	Mulher branca	Homem não branco	Mulher não branca
Sexo e raça	Pontuação média por sexo e raça	528,04 (0,30)	515,87 (0,25)	502,35 (0,15)	491,43 (0,11)
	Baixa renda	507,44 (0,31)	497,35 (0,24)	493,49 (0,15)	484,71 (0,11)
Renda	Não baixa renda	590,47 (0,60)	593,02 (0,56)	570,03 (0,46)	571,42 (0,46)
Tipo de escola	Escola pública e baixa renda	491,65 (0,69)	485,48 (0,52)	478,67 (0,3)	474,9 (0,22)
	Escola pública e não baixa renda	550,39 (2,66)	550,83 (2,78)	537,29 (1,58)	542,31 (1,65)
	Escola privada e baixa renda	565,22 (1,62)	562,74 (1,35)	549,73 (1,09)	548,91 (0,87)
	Escola privada e não baixa renda	622,36 (1,16)	621,53 (1,32)	603,06 (1,12)	602,71 (1,25)

Fonte: Elaboração dos autores a partir dos dados da pesquisa.

Nota: Os valores entre parênteses representam o desvio padrão de cada média.

Nas últimas linhas da tabela são consideradas as interações entre as variáveis consideradas neste quadro. Isto permite visualizar a média das notas considerando outros elementos além da média simples da pontuação, como por exemplo a pontuação média de acordo com o sexo, raça, nível de renda e tipo de escola do candidato.

Percebe-se que existe um padrão de heterogeneidade nas pontuações médias de acordo com o sexo e raça do candidato. O homem branco apresenta a maior média de nota, que é maior que a média da mulher branca em aproximadamente 12 pontos. A nota média do homem não branco é aproximadamente 11 pontos maior que da mulher não branca. Considerando a raça, a diferença entre as notas apresentadas na Tabela 4 se reverte, tornando a média de pontuação dos homens mais alta. Mas também reforça o resultado de desigualdade de desempenho apresentado na tabela anterior, de modo geral, as notas médias dos candidatos não brancos são mais baixas.

Considerando o nível de renda e o tipo de escola, o padrão desigual de desempenho mantém a mesma ordem apresentada ao se analisar apenas sexo e raça. O fato de os candidatos ser de escola pública também parece influenciar negativamente a pontuação média. Os candidatos que estudaram em escolas privadas, mesmo tendo baixa renda, apresentam uma média maior de notas. O fato de estudar em escola pública e ter baixa renda diminui ainda mais as médias de pontuação.

Deste modo, os alunos brancos com maior renda e que estudaram em escola pri-

vada apresentam a maior média de pontuação, enquanto que as mulheres não brancas com baixa renda e que estudaram em escolas públicas apresentam a média de pontuação mais baixa. Esses dados podem representar a desigualdade educacional e de oportunidades apresentada na literatura. Principalmente na região Nordeste, como apontado por Netto e Figueiredo (2010). O procedimento econométrico utilizado a seguir permite verificar mais detalhadamente a influência desses fatores sobre tais resultados

4.2 Testes

Esta pesquisa aborda como os fatores econômicos e sociais afetam a média das notas do ENEM no Nordeste. Antes de prosseguir com as estimações dos modelos, são realizados os testes do multiplicador de Lagrange, de Breusch-Godfrey/Wooldridge, de Jarque-Bera e de Breusch-Pagan aplicados nos resíduos dos modelos estimados. Os resultados desses testes são apresentados na tabela 5.

Tabela 5. Testes nos resíduos da regressão

Descrição Testes	Efeitos Fixos		Efeitos Aleatórios		Pooled	
	Estatística	p valor	Estatística	p valor	Estatística	p valor
Multiplicador de Lagrange	124826	0.000	124826	0.000	124826	0.000
Breusch-Godfrey/Wooldridge	32265	0.000	149.77	0.000	185.45	0.000
Jarque-Bera	48.307	0.000	2039	0.000	2037	0.000
Breusch-Pagan	3322.3	0.000	3322.3	0.000	3322.3	0.000

Fonte: Elaboração dos autores a partir dos dados da pesquisa.

O teste do multiplicador de Lagrange tem como hipótese nula a indicação de que todos os parâmetros β são iguais a zero em detrimento de que pelos menos algum dos estimadores do modelo é estatisticamente diferente de zero como hipótese alternativa. Então, observou-se que os modelos apresentaram, no mínimo, um dos estimadores estatisticamente diferentes de zero. Outro teste aplicado ao modelo foi o de correlação serial de Breusch-Godfrey/Wooldridge, que testa se a série tem ou não essa característica. A partir dos resultados obtidos, podemos notar que os resíduos dos modelos estimados apresentaram correlação serial.

Além dos testes citados, temos os de Jarque-Bera e o de Breusch-Pagan, que apontaram, respectivamente, que os resíduos dos modelos não apresentaram distribuição normal e são heterocedásticos (Tabela 6). Ambos rejeitaram as hipóteses nulas de normalidade dos resíduos e homoscedasticidade. Nesse sentido, os modelos estimados utilizaram os erros- padrões robustos de White “hc0”, que, segundo Souza (2011), é um dos mais precisos e utilizados quando se trata de modelos com erros robustos de heterocedasticidade.

ntes de escolher o agrupamento da série de dados em painel que utilizaremos, obtivemos os testes F de Chow para escolher entre o modelo de efeitos fixos e o *pooled*, o de Breusch-Pagan, para efeitos aleatórios ou *pooled*, e o de Hausman, para efeitos

Tabela 6. Testes de agrupamento dos dados

Testes	Estatística	p valor
F de Chow	10.145	0.020
Breusch-Pagan	47.143	0.000
Hausman	63.491	0.000

Fonte: Elaboração dos autores a partir dos dados da pesquisa.

fixos ou aleatórios. Na tabela 6, vemos que o teste F de Chow apontou que o modelo de efeitos fixos é preferível ao *pooled*; o de Breusch-Pagan indicou que o de efeitos aleatórios é mais apropriado que o *pooled*; e o teste de Hausman sugere o modelo de efeitos fixos ao invés de efeitos aleatórios.

5. Evidências Empíricas

Depois de aplicados os testes, na Tabela 7, apresentam-se as estimações dos modelos de dados em painel com erros robustos a heterocedasticidade de White por meio de efeitos fixos, aleatórios e *pooled*. Ressalta-se que tanto por meio do teste de Hausman quanto do critério de informação de Akaike (AIC), que considera como preferível o modelo com o menor valor do teste, o de efeitos fixos foi o indicado. Nesse sentido, as análises subsequentes são baseadas nesse modelo.

Na tabela 7, analisamos, por meio do método de dados em painel com efeitos fixos, os fatores econômicos e sociais que afetam o desempenho dos candidatos na região Nordeste no exame do ENEM no período de 2015 a 2019. De forma geral, todos os estimadores foram significativos estatisticamente de forma individual ao nível de significância de 1%, e todos resultados encontrados vão de encontro com o esperado pela literatura.

Em primeiro lugar, analisando o efeito da idade dos estudantes considerados, avalia-se que a idade afeta negativamente a média do ENEM, em que um ano adicional na idade do aluno, reduz a sua nota em média em 14,55 pontos, ou seja, ter menor idade e estar na série adequada para a idade beneficia o aluno que realiza o ENEM. Tal resultado corrobora as evidências obtidas por Silva (2017), que explicam que os alunos com idades mais avançadas passaram por evasão escolar e eventuais reprovações na escola, e isso afeta o desempenho do aluno no referido exame. Desse modo, quanto menor a distorção idade-série maior tende a ser a nota do aluno.

Em relação à idade², que apresenta um coeficiente positivo de 0,19, entende-se que esse grupo pode pertencer a participantes que desejam trocar de curso ou efetuar uma segunda graduação e que, segundo Torres et al. (2020), têm mais experiência e mais destreza para conseguir melhor rendimento no exame.

Quanto ao sexo dos participantes, o estudo revelou que os representantes masculinos apresentaram desempenho superior ao feminino. Especificamente, o coeficiente igual a -3,297 indica que ser mulher diminui a nota do ENEM em aproximadamente

Tabela 7. Estimacões dos modelos

	Efeitos Fixos	Efeitos Aleat3rios	Pooled
Vari3vel Dependente: M3dia no Enem			
Intercepto		700,626*** (4,382)	700,716*** (4,380)
idade	-14,551*** (0.583)	-14,805*** (0.334)	-14,811*** (0.334)
idade ²	0.190*** (0.011)	0.196*** (0.006)	0.196*** (0.006)
sexo	-3.297*** (0.559)	-1.644*** (0.330)	-1.637*** (0.330)
branco	7.374*** (0.661)	7.430*** (0.396)	7.434*** (0.396)
privado	50.267*** (0.943)	48.862*** (0.576)	48.854*** (0.576)
escpai	11.853*** (0.702)	12.909*** (0.412)	12.917*** (0.412)
escmae	12.394*** (0.660)	11.878*** (0.385)	11.878*** (0.385)
baixarenda	-40.378*** (1.095)	-41.062*** (0.685)	-41.062*** (0.685)
computador	17.956*** (0,055)	17.115*** (0,000)	17.117*** (0,000)
internet	7.842*** (0.681)	10.695*** (0.393)	10.715*** (0.393)
moradores	-1.714*** (0.191)	-2.126*** (0.108)	-2.128*** (0.109)
urbana	5.837*** (1.387)	7.420*** (0.776)	7.428*** (0.776)
celular	13.760*** (1.409)	10.978*** (0.812)	10.958*** (0.812)
BIC	1869938	2069215	2070937
AIC	1869796	2069063	2070785

Fonte: Elaborac3o dos autores a partir dos dados da pesquisa.

3,3 pontos se comparado a ser homem. Isso ratifica os resultados encontrados por Menezes e Soares (2010) e Torres et al. (2020) os quais observaram que os homens, em m3dia, apresentam melhores desempenhos em Matem3tica quando comparados com as mulheres em todos os n3veis de escolaridade. Por sua vez, Silva (2017), encontrou tais evidencias nas 3reas de Ci3ncias Humanas, Matem3tica e Ci3ncias da Natureza.

O modelo tamb3m indica que os candidatos de cor branca exerceram desempenho superior ao das demais raas, assim como os trabalhos de Figueiredo et al. (2014); Torres et al. (2020) e Silva (2017). Conforme Andrade (2012), isso pode ocorrer porque a cor do aluno, geralmente, demonstra sua herana cultural, o que pode explicar seu desempenho. Al3m disso, ocorre um cen3rio em que os sal3rios das pessoas n3o brancas s3o menores do que os das brancas. Isso esclarece que o desempenho superior dos alunos brancos em rela3o aos pardos e aos pretos, pode se dar devido 3s condi3es socioecon3micas superiores da popula3o branca.

O tipo de ensino tamb3m implicou o rendimento do ENEM, indicando que os estudantes de escolas privadas tiveram um desempenho superior em 50,27 pontos aos de

escolas públicas. Este resultado corrobora com Menezes-Filho (2007) ao constatarem um desempenho melhor de alunos de escolas privadas se comparados com alunos de escola pública. Ao encontrarem resultados semelhantes, Figueiredo et al. (2014) afirmam que esse resultado está associado a diferença de qualidades entre as escolas públicas e privadas, que é uma das principais causas da desigualdade de oportunidade educacional do Brasil.

Quanto à escolaridade da mãe e do pai do candidato, verificou-se que estas afetaram positivamente o rendimento, o que se confirma nos trabalhos de Figueiredo et al. (2014), Mendes e Karruz (2012) e Torres et al. (2020), que enunciam que, quanto maior o nível de educação dos pais, maiores são os estímulos que eles dão aos filhos para estudarem.

O fato de o candidato pertencer à classe econômica de baixa renda familiar repercutiu negativamente no rendimento no ENEM, como citado por Figueiredo et al. (2014), Torres et al. (2020) e Silva (2017). Isso se justifica porque, se tiverem uma renda maior, os candidatos podem adquirir melhores materiais de estudo e cursos. Assim como no trabalho de Figueiredo et al. (2014), a localização da residência do aluno também afetou o desempenho no ENEM, pois os estudantes residentes na zona urbana têm médias mais altas do que os que moram na zona rural. A quantidade de moradores na residência do candidato contribui negativamente para o desempenho no ENEM, como demonstrado no trabalho de Lobo et al. (2017). A presença de celulares na residência afetou positivamente o rendimento no ENEM. De acordo com Araújo (2019), com mais de 80% de confiança, os alunos que obtiveram acima de 750 pontos no ENEM disseram que têm, em casa, três celulares ou mais, no período de 2016 a 2018.

A presença do computador afetou positivamente o rendimento no ENEM, o que corrobora o trabalho de Lobo et al. (2017) assim como os de Torres et al. (2020) e Silva (2017), segundo os quais o acesso à internet contribuiu para elevar as notas no ENEM. A junção do computador com o celular e o acesso à internet contribui para difundir informações e possibilita aos alunos acessarem uma infinidade de materiais de estudo, que, se bem e explorados, podem melhorar o rendimento no ENEM. Portanto, os sinais positivos nos parâmetros podem ser explicados a partir dessa relação.

Por fim, observa-se que o fato da região Nordeste possuir 77,7% dos candidatos não brancos e 88,0% de baixa renda, aliadas as desigualdades raciais e socioeconômicas, possivelmente corrobora para o baixo desempenho relativo desta região em relação as demais brasileiras. Tendo em vista que como apontado pela Tabela 7, ser não branco e de baixa renda contribui em média para uma pontuação baixa no ENEM.

6. Considerações finais

A literatura aponta que os diversos fatores econômicos e sociais como idade, sexo, raça, tipo de ensino, localização da escola, escolaridade dos pais, quantidade de mo-

radores na residência, se percebem baixa renda e se têm acesso a equipamentos eletrônicos que possibilitem o acesso à informação afetam o desempenho dos candidatos no exame do ENEM. Nesse sentido, esta pesquisa buscou analisar como o desempenho dos candidatos no ENEM, na Região Nordeste do Brasil, pode ser afetado por meio dos fatores ligados às características sociais e econômicas locais.

Os resultados alcançados sugerem que principalmente a escolaridade dos pais, o tipo da escola e a disposição de ferramentas como celular, computador e acesso à internet contribuem para aumentar a média dos candidatos situados na Região Nordeste. Por outro lado, destacam-se como limitantes das médias a idade dos candidatos e se pertencem à classe econômica com baixa renda. Provavelmente, a idade do aluno indica que ele sofreu alguma evasão escolar, o que dificulta seu rendimento no exame. E o fato de pertencer a uma classe com baixa renda reduz as possibilidades de melhores condições de aprendizagem, como cursos e materiais de estudo, por exemplo.

Nesse sentido, este estudo traz uma nova contribuição para a literatura sobre avaliação de desempenho escolar ao realizar um corte regional para o Nordeste, tendo sido analisado um período recente. Como indicam os dados educacionais e a literatura esta é uma região que vem sofrendo com a desigualdade educacional, além de outros problemas socioeconômicos que afetam o desempenho dos alunos durante a jornada escolar e em provas como o ENEM.

Portanto, os fatores propulsores da elevação da média no ENEM devem ser observados e estimulados pelos governantes locais, como a concessão, por parte do governo, de celulares e/ou computadores com acesso à internet para os alunos com reduzida capacidade financeira, que, além de reduzir as disparidades entre alunos pobres e ricos, eleva a média geral do exame e promove ganhos para a sociedade de forma geral.

Em relação às dificuldades de desenvolver a pesquisa, destaca-se a metodologia da base de dados do ENEM, que, entre os anos de 2009 e 2014, passou por várias mudanças, que dificultaram analisar um período maior. Como proposta de pesquisas futuras, uma alternativa é a utilização de uma base de dados que contemple o período de 2009 a 2019, a partir do método de dados em painel, visando captar efeitos no tempo e no espaço.

Referências

- Albernaz, A., Ferreira, F., e Franco, C. (2002). Qualidade e equidade no ensino fundamental brasileiro. *Pesquisa e Planejamento Econômico*, 32.
- Andrade, C. Y. d. (2012). Acesso ao ensino superior no Brasil: equidade e desigualdade social. *Revista Ensino Superior*, 6:18–27.

- Araújo, D. L. d. (2019). Determinantes do desempenho no ENEM dos concluintes do ensino médio no município de Viçosa - MG.
- Baltagi, B. H. (2005). *Econometric analysis of panel data*. Chichester: John Wiley and Sons, 3:302.
- Cunha, F., J.HECKMAN, J., LOCHNER, L., e V.MASTEROV, D. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1:697–812.
- Di Pietra, G., Kawakoa Komatsu, B., Sasaki, A., e Menezes-Filho, N. (2018). Por que o Brasil vai Mal no PISA? Uma Análise dos Determinantes do Desempenho no Exame. *INSAPER – Centro de Políticas Públicas*.
- Figueiredo, E., Nogueira, L., e Santana, F. L. (2014). Igualdade de Oportunidades: Analisando o papel das circunstâncias no desempenho do ENEM. *Revista Brasileira de Economia*, 68:373–392.
- Greene, W. H. (2007). Fixed and Random Effects Models for Count Data. *NYU Working Paper No.*, Página 07–16.
- Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions . *Journal of Human Resources*.
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46:1251–1271.
- INEP (2019). Dados do ENEM.
- Jesus, G. R. d. e Laros, J. A. (2004). Eficácia escolar: regressão multi-nível com dados de avaliação em larga escala. *Avaliação Psicológica*, 3:93–106.
- Levacic, R., Jenkins, A., Vignoles, A., Steele, F., e Allen, R. (2005). Estimating the relationship between school resources and pupil attainment at Key Stage 3. *Department for Education and Skills/Institute of Education, University of London, Londres*.
- Lobo, G. D., Cassuce, F. C. D. C., e Cirino, J. F. (2017). Avaliação do desempenho escolar dos estudantes da região nordeste que realizaram o ENEM: uma análise com modelos hierárquicos. *Revista Espacios*, 6.
- Macedo, G. A. (2004). Fatores associados ao rendimento escolar de alunos da 5ª série (2000): uma abordagem do valor adicionado. *Trabalho apresentado no XIV Encontro Nacional de Estudos Populacionais, ABEP*.
- Marques, L. D. (2000). Modelos dinâmicos com dados em painel: revisão de literatura. *Centro de estudos Macroeconómicos e Previsão, faculdade de Economia do Porto*.
- Melo, L. M. C. e Arakawa, V. H. (2012). Existe desigualdade regional na relação entre background familiar e desempenho escolar dos filhos? Evidências para as grandes regiões do Brasil. *XVIII Encontro Nacional de Estudos Populacionais, ABEP*.

- Mendes, B. D. e Karruz, A. P. (2012). Background familiar, desigualdade regional e o desempenho no Exame Nacional do Ensino Médio (ENEM). *Universidade Federal de Minas Gerais (UFMG)*, Página 1–25.
- Menezes, T. e Soares, S. (2010). Os determinantes do desempenho escolar: uma análise para o estado de pernambuco. *Recife, Fórum BNB*.
- Menezes-Filho, N. A. (2007). Os determinantes do desempenho escolar do Brasil. *Instituto Futuro Brasil/IBMEC*.
- Montalvão, A. L. e Neubert, L. F. (2016). Desigualdades raciais e desempenho acadêmico no Brasil- Relações raciais: desigualdades, identidades e políticas públicas. *40º Encontro Anual da Anpocs*.
- Netto, J. L. S. e Figueiredo, E. A. (2010). Distribuição de capital humano e desigualdade de renda: Mobilidade Intergeracional Educacional e Mobilidade de Renda no Brasil. *Revista Economia e Desenvolvimento*, 8.
- Queiroz, L. O., Cassuce, F. C. C., e Cirino, J. F. (2018). Os efeitos da demissão do chefe de família sobre o acesso dos seus dependentes ao ensino superior. *Planejamento e Políticas Públicas*, 1:95–123.
- Rodrigues, L. d. O., Araujo, J. A., Guedes, J. P. M., e Silva, M. M. d. C. (2017). Mensuração da desigualdade educacional entre os municípios nordestinos. *Revista de Economia Contemporânea*, 21.
- Silva, I. C. d. S. (2017). Fatores que influenciam o desempenho dos participantes da prova ENEM: um estudo para os residentes do Ceará no ano de 2015. *Universidade Federal do Ceará*.
- Soares, J. F. (2004). O efeito da escola no desempenho cognitivo de seus alunos. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 2.
- Souza, T. C. d. (2011). Ensaio sobre modelos de regressão com dispersão variável. *Universidade Federal de Pernambuco*.
- Steele, F., Vignoles, A., e Jenkins, A. (2007). The effect of school resources on pupil attainment: a multilevel simultaneous equation modelling approach. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 170:801–824.
- Torres, R., Pereira, M. M., Filho, R. B., e Lisbinski, F. C. (2020). Determinantes do Desempenho dos Participantes da Prova do Enem: Evidências para o Rio Grande Do Sul. *Desenvolvimento em Questão*, 18:352–368.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.