

Avaliação do Desempenho Escolar dos Estudantes da Região Nordeste que Realizaram o ENEM: Uma Análise com Modelos Hierárquicos.

Área 2: Economia

Gustavo Dantas Lobo
Universidade Federal de Viçosa - UFV
Departamento de Economia
gustavo.lobos@ufv.br

Francisco Carlos da Cunha Cassuce
Universidade Federal de Viçosa - UFV
Departamento de Economia
francisco.cassuce@ufv.br

Jader Fernandes Cirino
Universidade Federal de Viçosa - UFV
Departamento de Economia
jader.cirino@ufv.br

Resumo:

Neste trabalho construiu-se um modelo hierárquico em dois níveis, compreendendo tanto variáveis referentes ao *background* familiar do aluno (1º nível) quanto do tipo de escola (2º nível), a fim de identificar os fatores determinantes do desempenho escolar e qual a proporção explicada da variância total das notas por cada agrupamento. Utilizaram-se os microdados do ENEM para alunos nordestinos do 3º ano do ensino médio na prova de matemática em 2013. Verificou-se que 15.28% da variância total das notas é explicada pelas diferenças entre escolas e variáveis como a escolaridade dos pais e nível de renda, tiveram impactos positivos no desempenho.

Palavras-chave: Educação, Modelos Hierárquicos, ENEM, Nordeste.

I.INTRODUÇÃO

A educação pode ser considerada como uma das ferramentas mais eficientes de se romper com o ciclo da pobreza. Sendo assim, a busca pelo desenvolvimento educacional é um dos principais objetivos a serem perseguidos pelos países, e com o Brasil não vem sendo diferente.

A partir da década de 1990, no que diz respeito à universalização do ensino, o Brasil conseguiu certo êxito, visto que no início da década, a taxa de frequência entre crianças de 7 a 14 anos era de 84,43%, alcançando 98,70% em 2014. A média de anos de estudo, por sua vez, passou de 4,8 anos em 1990 para 7,8 em 2014, um aumento substancial (IPEA, 2015).

No que se refere aos gastos totais do governo em educação em relação ao PIB, o país investiu cerca de 6,2% do PIB no ano de 2013, valor este que já chegou a ser bem menor, 4,6% no ano 2000. Ao se tratar de investimento direto em educação por aluno, o Brasil investiu cerca de R\$6.203,00 no ano de 2013, montante este que já foi 65% menor, representando um total de R\$2.197,00 por aluno em 2000 (INEP, 2013).

Mesmo apresentando uma evolução no que diz respeito aos gastos com educação e à universalização, o país ainda não conseguiu colher os frutos em termos de proficiência nos resultados, tanto nos exames nacionais quanto nos internacionais. A prova Brasil, exame destinado à avaliação do aprendizado dos alunos do ensino básico, constatou que apenas 23% dos alunos do nono ano da rede pública aprenderam o adequado na competência de leitura e interpretação de texto. Da mesma forma, apenas 11% destes mesmos alunos aprenderam o adequado na competência de resolução de problemas matemáticos (INEP, 2013).

Nos exames internacionais como o Pisa, os resultados também não são satisfatórios. Apesar da média alcançada nas notas de matemática ser a que mais evoluiu no mundo desde 2003 (de 356 para 391 em 2012), o Brasil ainda opera abaixo da média dos países da OCDE, tendo resultados inferiores a países como Chile, Costa Rica, México e Uruguai. Nas provas de português e ciências, o Brasil também apresenta resultados abaixo da média mundial, mesmo tendo apresentado melhoras nos últimos anos. Dos 65 países associados à OCDE, o Brasil ocupa a 58ª, superando países latino-americanos como Argentina, Colômbia e Peru (OCDE, 2012).

Com o intuito de direcionar políticas públicas, os economistas vêm tentando compreender e identificar os fatores determinantes para um bom desempenho escolar. A tendência apontada pelos trabalhos tanto nacionais quanto internacionais, é a de que as características referentes à condição socioeconômica do indivíduo se sobrepõem às características da escola e dos professores quando se trata de rendimento escolar. No caso brasileiro, Menezes-Filho (2007) verificou através dos dados do SAEB, quais os determinantes do desempenho escolar. Para tal, o autor utilizou variáveis referentes aos alunos, à escola e aos professores, chegando à conclusão de que são as características familiares e do aluno que mais explicam o desempenho escolar dos alunos, enquanto as variáveis referentes aos professores e às escolas têm efeitos reduzidos sobre o desempenho, chegando a resultados semelhantes aos que são verificados em todo o mundo.

Visando compreender a estratificação educacional, Riani e Rios-Neto (2008) utilizaram-se de um modelo logístico hierárquico a fim de identificar os determinantes de um indivíduo estar cursando o ensino formal na idade correta ou não. Para tal, foram utilizadas variáveis referentes ao *background* familiar e da qualidade do ensino. A principal conclusão identificada pelos autores foi que o nível educacional da mãe foi a variável que mais influenciou na probabilidade de um indivíduo estar cursando o ensino

formal, porém, constatou-se a existência de um efeito substituição entre educação materna e os fatores de qualidade da rede escolar dos municípios. Ou seja, fatores que influenciam na qualidade do ensino ofertado são capazes de minimizar as diferenças causadas pelo menor nível educacional materno.

Visto isso, este trabalho pretende identificar qual o papel do tipo da escola (pública ou privada) no desempenho dos alunos do 3º ano do ensino médio da região Nordeste na prova de matemática do Exame Nacional do Ensino Médio (ENEM). Para tal, estima-se um modelo hierárquico em dois níveis de intercepto aleatório em que o primeiro nível se refere à situação socioeconômica do aluno, e o segundo referente à situação administrativa da escola. Portanto, além de calcular o impacto do tipo de escola na variância total do desempenho escolar, estima-se o impacto das variáveis referentes ao aluno nas notas do ENEM.

A escolha da região Nordeste como objeto de estudo se deu devido ao seu histórico atraso socioeconômico. Mesmo com uma visível evolução em seus indicadores, a região ainda apresenta uma carência profunda, o que implica em uma maior necessidade de se analisar e se identificar possíveis alternativas ao desenvolvimento da região. Detecta-se que 7.43% dos domicílios Nordestinos em 2014 possuíam renda domiciliar per capita inferior à linha de extrema pobreza. Tal população equivale à quase 56% de toda a população extremamente pobre do Brasil. A região possui o pior índice de Gini dentre todas as regiões brasileiras, 0.516, valor este ligeiramente inferior ao índice nacional, que é 0.518. A região ainda apresenta a menor renda média mensal dentre todas as regiões. Enquanto a renda média brasileira em 2014 foi de R\$1,152.24, a renda média nordestina foi de apenas R\$730.24, uma renda 37% menor que a média nacional (IPEADATA).

No que diz respeito aos indicadores educacionais, a região Nordeste possui 3.48% de suas crianças de 10 a 14 anos analfabetas, constituindo a maior proporção dentre todas as regiões brasileiras. Vale ressaltar, que a proporção de crianças analfabetas de todo o Brasil é de 1.78%. Além disso, a região possui a menor média de anos de estudo entre as pessoas de 25 anos ou mais, 6.4 anos de estudo, valor este bem inferior à média nacional, que é de 7.8 anos de estudo (IPEADATA). A região também apresentou a menor taxa de aprovação dentre todas as regiões em 2014, 83.2%, enquanto a taxa nacional foi de 89.2%. Em relação ao nível de instrução dos docentes, 88.6% dos professores nordestinos do ensino médio possuíam educação superior em 2014, sendo mais uma vez, a menor taxa dentre todas as regiões, inclusive menor que a taxa brasileira, que foi 93.2% (INEP, 2014).

Visando corroborar com os trabalhos que tratam de características dos indivíduos como determinantes do desempenho escolar, pretende-se com este artigo explorar quais os efeitos de variáveis como a escolaridade dos pais, renda, cor, sexo, tamanho da família, localização do domicílio, acesso à internet e se o indivíduo trabalha ou não. A estimação será feita por regressão multinível, tentando identificar os impactos no desempenho estudantil entre aqueles que cursaram o ensino público e o privado.

II. REVISÃO DE LITERATURA

No final dos anos 1950, o tema da educação passou a ser recorrente na pesquisa econômica. Trabalhos seminais como os de Schultz (1961), Becker (1962) e Mincer (1958), deram um caráter de decisão de investimento para a educação, em que o indivíduo incorre em custos (pecuniários e de oportunidade), a fim de obter um aumento de renda no futuro propiciado pela maior escolaridade, caracterizando-se então como o investimento em capital humano.

Mincer (1958) por sua vez trabalhou em uma forma de mensurar os retornos da educação, estimando o impacto de um ano a mais de educação sobre os salários através de um coeficiente de educação.

A partir da perspectiva de Mincer, muitos autores se dedicaram à estimar os retornos da educação no Brasil, destacam-se, Ueda e Hoffmann (2002), Sachsida et al. (2004) e Resende e Wyllie (2006). Em geral, a literatura reporta taxas de retorno entre 10% e 20%, sendo maiores na cidade do que no meio rural (Barbosa Filho e Pessôa, 2013).

Em uma perspectiva voltada para o crescimento econômico, os modelos de crescimento endógeno mostram que elevações do nível de educação causam elevação permanente da taxa de crescimento da renda per capita. Entre eles, o mais tradicional é o de Solow (1956) aumentado, em que os investimentos em capital humano aumentam a produtividade total dos fatores empregados na economia. Outros trabalhos, como os de Lange e Topel (2006) e Card (1999) tentam mostrar que há causalidade entre nível educacional e produtividade.

Quanto a geração e desigualdade de renda, Barros e Mendonça (1995) mostraram que os níveis desiguais de educação são responsáveis por 35% a 50% da elevação da desigualdade de renda.

Nesse sentido, o nível e a qualidade educacional oferecidos em um país passam a ser fundamentais não só para a geração de atividade econômica, mas também para a geração e distribuição de renda. Diante disso, ficaria explícita a importância de se analisar os principais determinantes capazes de aumentarem o rendimento dos estudantes e melhorarem a qualidade da educação brasileira.

Existe um extenso debate na literatura de economia da educação a respeito dos determinantes do desempenho escolar. Fatores ligados ao background familiar do aluno se mostraram, tanto na literatura internacional quanto na nacional, os mais impactantes em termos de rendimento. São eles: renda familiar, escolaridade dos pais, principalmente da mãe, o número de membros no domicílio do aluno, o trabalho na infância, o acesso às tecnologias da informação como computador e internet, a leitura de livros e jornais. Mesmo os fatores referentes à escola e aos professores sendo menos impactantes, eles também tem sua parcela de responsabilidade no desempenho escolar. Fatores como a estrutura física da escola, presença de laboratórios, o tamanho das turmas, o salário do professor e o seu nível de instrução são as variáveis mais frequentemente estudadas.

O debate a cerca dos determinantes do desempenho escolar tomou grandes proporções a partir do relatório Coleman, liderado principalmente por Eric Hanushek, que rechaçou o argumento até então vigente de que o volume de recursos destinados à educação era o principal fator que proporcionava melhores resultados escolares. Hanushek argumenta que os fatores relacionados ao *background* familiar têm maior relevância no desempenho dos alunos. (HANUSHEK, 1989).

No Brasil, inúmeros autores também confirmaram os resultados encontrados pelo relatório Coleman, com destaque para Menezes-Filho (2007). Em seu trabalho, foram levadas em consideração variáveis referentes aos alunos e suas famílias, à escola e aos professores, chegando a conclusão de que as variáveis referentes à situação socioeconômica do aluno são as preponderantes no desempenho do mesmo. Um fator fundamental apontado pelo autor é o fato de alunos brancos terem melhores desempenhos que os negros. Isto pode se dar por características não observáveis como a escolaridade dos pais e a renda, que proporcionam um menor aprendizado por parte dos negros.

Tentando determinar os principais fatores que influenciam no desempenho em matemática de alunos da rede pública estadual de Minas Gerais, Machado *et al* (2008), através de um modelo hierárquico de três níveis, estabeleceu que fatores relacionados ao aluno como a escolaridade da mãe e a presença de livros em casa são determinantes no desempenho escolar.

A taxa de fecundidade e a sua relação com a educação foi objeto de estudo de Berquó e Cavenaghi (2004) e Lam e Duryea (1999), que mostram que há um elevado diferencial de fecundidade associado à escolaridade da mãe, demonstrando a importância da educação como forma de reduzir a taxa de natalidade. Mães com maior nível de escolarização tem maior capacidade de planejar e gerir a decisão de ter filhos e cuidar de uma família, podendo prezar, portanto, pela educação de seus próprios filhos.

Seguindo a mesma linha, Barros *et al* (2001) identificou a escolaridade dos pais, em particular a da mãe, como o principal fator que afeta o desempenho estudantil, superando fatores como a renda per capita familiar e o nível de instrução dos professores. Fatores como a proximidade da escola em relação ao domicílio do aluno, presença de livros e laboratórios também são estatisticamente significativos, já que reduzem os custos de transação de se estudar. Além disso, regiões que apresentam altos custos de oportunidade do tempo, ou seja, aquelas que possuem maiores oportunidades no mercado de trabalho, tendem a possuir pior desempenho nos indicadores educacionais.

Albernaz, Ferreira e Franco (2002) encontraram, através de modelos hierárquicos lineares, que cerca de 80% da variância em desempenho médio entre as escolas deve-se às diferenças socioeconômicas dos estudantes. Além disso, fatores como o fato de a escola ser pública ou privada, a escolaridade dos professores e variáveis referentes à infraestrutura física da escola também são fundamentais para explicar o desempenho estudantil.

De forma análoga, Jesus e Laros (2004), encontraram que 79% da variância entre escolas no desempenho em língua portuguesa são atribuídas principalmente à variáveis relacionadas à condição socioeconômica tanto dos alunos quanto das escolas.

Biondi e Felício (2007) tentaram identificar quais fatores referentes à escola que determinam o desempenho médio dos alunos. Entre os efeitos significativos encontrados neste estudo, tem-se que a ausência de rotatividade dos professores ao longo do ano letivo, a experiência média dos professores superior a dois anos em sala de aula e a existência na escola de conexão com a Internet afeta positivamente o resultado médio. Em contrapartida, a existência na escola de laboratório de informática está negativamente relacionada com o desempenho.

Andrade e Laros (2007), utilizando-se de regressões multinível tanto para alunos quanto para escolas, encontraram uma correlação intraclasse de 0,17, ou seja, 17% da variância no desempenho escolar podem ser atribuídos ao nível da escola. Além disso, variáveis como o atraso escolar e a comparação do aluno com os colegas são as variáveis referentes aos alunos que mais afetam o desempenho escolar.

III. METODOLOGIA

Utilizou-se neste trabalho os microdados referentes ao Exame Nacional do Ensino Médio (ENEM) de 2013 para os alunos da região Nordeste. Foram escolhidos aqueles alunos que estavam cursando o 3º ano do ensino médio e que realizaram o exame com o intuito de ingressar em alguma instituição de ensino. No total, definiu-se uma amostragem de 329,673 indivíduos. Destes, 21.78% estão distribuídos entre as escolas particulares, enquanto 78.22% dos demais alunos estão inseridos na rede pública

de ensino. Na tabela 1 à seguir, são definidas todas as variáveis utilizadas e qual o seu efeito esperado sobre o desempenho escolar em matemática.

Tabela 1: variáveis utilizadas no modelo.

Variáveis	Descrição
Nota em Matemática	Variável explicativa. Nota alcançada no exame de matemática.
Nº de membros do domicílio	Número de componentes do domicílio. Espera-se um efeito negativo do tamanho do domicílio na nota média.
Sexo	Sexo do estudante. Assume valor 1 para masculino e 0 para feminino. Espera-se um efeito positivo no desempenho escolar para aqueles do sexo masculino.
Cor	Cor do aluno. Assume o valor 1 para branco e 0 para não branco. Espera-se um efeito positivo no desempenho em matemática para aqueles alunos brancos.
Classe A	Nível de renda familiar acima de 20 salários mínimos. Assume valor 1 caso pertencente, e 0 caso contrário. Para este nível de renda, espera-se um efeito alto e positivo no desempenho escolar.
Classe B	Nível de renda familiar de 10 a 20 salários mínimos. Assume valor 1 caso pertencente, e 0 caso contrário. Para este nível de renda, espera-se um efeito positivo no desempenho escolar.
Classe C	Nível de renda familiar de 4 a 10 salários mínimos. Assume valor 1 caso pertencente, e 0 caso contrário. Para este nível de renda, espera-se um efeito positivo, porém modesto no desempenho escolar.
Classe E	Nível de renda familiar de 2 salários mínimos. Assume valor 1 caso pertencente, e 0 caso contrário. Para o nível mais baixo de renda, espera-se um efeito negativo sobre o desempenho escolar.
Instrução fundamental da mãe e/ou do pai	Nível educacional mais alto alcançado na família. Assume o valor 1 quando o maior nível é o fundamental e 0 quando não. Assume-se que quanto mais baixo o nível de instrução dos pais, menor é o efeito no desempenho escolar.
Instrução média da mãe e/ou do pai	Nível educacional mais alto alcançado na família. Assume o valor 1 quando o maior nível é o médio e 0 quando não. Assume-se que quanto mais baixo o nível de

	instrução dos pais, menor é o efeito no desempenho escolar.
Instrução superior da mãe e/ou do pai	Nível educacional mais alto alcançado na família. Assume o valor 1 quando o maior nível é o superior e 0 quando não. Assume-se que quanto mais baixo o nível de instrução dos pais, menor é o efeito no desempenho escolar.
Internet	Aponta a presença de internet no domicílio do aluno. Assume valor 1 para a presença e 0 caso contrário. Espera-se que seu efeito seja positivo sobre o desempenho escolar.
Computador	Aponta a presença de computador no domicílio do aluno. Assume valor 1 para a presença e 0 caso contrário. . Espera-se que seu efeito seja positivo sobre o desempenho escolar.
Trabalha	Indica se o aluno trabalha ou já trabalhou. Assume valor 1 para aqueles que trabalham ou trabalharam e 0 caso contrário. Presume-se que o tempo dedicado ao trabalho afete negativamente o desempenho escolar.
Escola Pública	Assume valor um para aqueles indivíduos que estudam em escola pública e 0 para aqueles que frequentam escola privada. Esta é a única variável referente ao nível da escola. Espera-se que haja um efeito negativo por se estudar em escola pública.

Fonte: microdados do ENEM 2013.

Os modelos multinível podem ser considerados extensões dos modelos de regressão linear tradicionais, em que se estimam regressões para níveis diferentes de análise. Isto se dá devido ao fato de os indivíduos poderem estar inseridos em diferentes contextos, portanto sofrem efeitos diferentes das variáveis explicativas.

Para Cruz (2010), os modelos multinível são uma resposta à necessidade de analisar as relações entre os indivíduos e o meio em que estão inseridos, podendo separar o papel de cada um das características de um contexto, com o intuito de melhorar o conhecimento da realidade, podendo, caso necessário, intervir de forma mais eficiente.

Neste trabalho, pretende-se analisar o efeito das variáveis socioeconômicas no desempenho escolar do aluno, mais especificamente na disciplina de matemática, porém a nível de escola, sendo ela pública ou privada. Ou seja, pretende-se analisar o efeito que o tipo de escola tem no desempenho do aluno. Para tal, estima-se um modelo em dois níveis com intercepto aleatório, o primeiro referente ao aluno, e o segundo, referente à escola:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \quad (1)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} \quad (3)$$

Em que Y_{ij} é o desempenho do estudante, β_{0j} é o intercepto, ou seja, o valor médio do desempenho dos alunos quando o conjunto de variáveis independentes X_{ij} for zero, β_{1j} é o coeficiente angular da regressão, X_{ij} é uma matriz de variáveis independentes referentes ao nível dos alunos, Z_j é a variável referente ao tipo da escola. Os termos de erro e_{ij} e u_{0j} referem-se aos níveis 1 e 2 respectivamente e são normalmente distribuídos com média 0 e variância σ^2 . Por fim, o termo γ_{10} representa a média dos coeficientes de regressão das j escolas para as variáveis independentes. O subscrito i é referente a cada indivíduo ($i=1, \dots, n_j$) e o subscrito j se refere à escola ($j=1, \dots, J$).

Percebe-se que o intercepto depende diretamente do tipo de escola. Ou seja, se γ_{00} e γ_{01} são positivos, eles afetam positivamente o intercepto, interferindo diretamente no desempenho médio escolar.

O termo de erro do segundo nível (u_{0j}) possui distribuição normal e são independentes do erro do primeiro nível (e_{ij}). Além disso, pode-se perceber que os coeficientes de regressão γ não variam de acordo com o tipo de escola, portanto são considerados como coeficientes de efeito fixo. Os efeitos variáveis entre as escolas são determinados pela variação dos erros dos resíduos de ambos os níveis. Isto pode ser melhor verificado no modelo extensivo, substituindo (2) e (3) em (1):

$$Y_{ij} = \gamma_{00} + \gamma_{01}Z_j + \gamma_{10}X_{1ij} + u_{0j} + e_{ij} \quad (4)$$

Os argumentos que possuem γ são considerados a parte fixa ou determinística do modelo, enquanto aqueles que possuem termos de erro são considerados a parte randômica ou estocástica do modelo.

A dependência das observações pode ser medida por um coeficiente de correlação: a correlação intra-classe. Ele pode ser definido pela seguinte equação:

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_e^2 + \sigma_{u_0}^2} \quad (5)$$

A correlação intra-classe indica a proporção da variância explicada pela estrutura de agrupamento na população, ou seja, o montante da variância total do desempenho dos alunos nordestinos que é explicada pelo tipo de escola. A equação nos mostra que a correlação intra-classe é a proporção da variância de um nível comparado com a variância total. Coeficientes muito baixos indicam que não há necessidade alguma de se realizar a análise em mais de um nível.

Para garantir a adequação do modelo, estimou-se um teste χ^2 , o LR *test*, que tem como hipótese nula o fato dos efeitos aleatórios serem iguais a zero. Este teste é realizado para confirmar a necessidade de se fazer a análise em diferentes níveis ou em um único nível.

Além disso, foi realizado um teste de razão de máxima verossimilhança a fim de comparar o modelo de intercepto aleatório com um modelo de intercepto e inclinação aleatórios. Para tal, estimou-se um modelo em que o nível de renda da família é função do tipo de escola que o aluno frequenta. A hipótese nula do teste é de que os modelos não apresentam diferenças significativas contra a hipótese alternativa de que eles são distintos.

Para mais detalhes à respeito desta metodologia, consultar da Cruz (2010) e Hox (2010).

IV. RESULTADOS

Primeiramente, com o intuito de caracterizar a região em estudo, são apresentadas na tabela 2 as estatísticas descritivas das variáveis definidas no modelo para a região Nordeste em contraposição aos da região Sudeste. A região Sudeste foi escolhida como base comparativa devido ao seu destaque em diversos indicadores socioeconômicos, sendo assim, a região mais desenvolvida do país.

Tabela 2: estatísticas descritivas das regiões Nordeste e Sudeste.

Variáveis	Média		Desvio padrão		Valor mínimo		Valor máximo	
	Nordeste	Sudeste	Nordeste	Sudeste	Nordeste	Sudeste	Nordeste	Sudeste
Nota Mat.	493.2375	536.9411	101.6064	106.7499	327.5	324.3	970.9	971.5
Nota Mat. Esc. Pub.	470.1138	509.8763	85.00839	91.87295	327.5	324.3	970.9	971.5
Nota Mat. Esc. Part.	576.2816	615.7882	112.3886	107.9499	327.5	327.5	970.9	971.5
Sexo	0.4054	0.4178	0.4909	0.4932	0	0	1	1
Cor	0.2633	0.5405	0.4404	0.4983	0	0	1	1
Membros Dom.	4.3978	4.0875	1.4996	1.2731	1	1	20	20
Classe A	0.0109	0.0224	0.1037	0.1481	0	0	1	1
Classe B	0.0202	0.0463	0.1407	0.2103	0	0	1	1
Classe C	0.0722	0.1832	0.2589	0.3868	0	0	1	1
Classe E	0.7857	0.4820	0.4102	0.4996	0	0	1	1
Trabalha	0.2117	0.3934	0.4085	0.4885	0	0	1	1
Internet	0.5541	0.8278	0.4970	0.3775	0	0	1	1
Comp.	0.5518	0.8197	0.49731	0.3843	0	0	1	1
Instrução sup. Mãe	0.1434	0.2101	0.3505	0.4074	0	0	1	1
Instrução méd. Mãe	0.2860	0.3336	0.4519	0.4715	0	0	1	1

Instrução o fun. Mãe	0.5172	0.4376	0.4997	0.4961	0	0	1	1
Instrução o sup. Pai	0.0903	0.1683	0.2865	0.3741	0	0	1	1
Instrução o méd. Pai	0.2394	0.3124	0.4267	0.4635	0	0	1	1
Instrução o fun. Pai	0.5753	0.4933	0.4942	0.4999	0	0	1	1
Escola Pública	0.7822	0.7444	0.4127	0.4361	0	0	1	1

Fonte: resultados da pesquisa.

Pode-se perceber, através dos dados apresentados, a discrepância existente entre as duas regiões. Primeiramente nota-se que o desempenho médio em matemática dos alunos do 3º ano do Nordeste é 43.70 pontos menor do que a média do Sudeste, cerca de 8%. Ao se analisar a diferença entre escolas públicas e privadas entre as regiões, verifica-se uma superioridade em termos de rendimento da região Sudeste em relação ao Nordeste para os dois tipos de escola. Além disso, fica evidente que o desempenho médio dos alunos de escolas particulares supera bastante o dos estudantes da rede pública de ensino, alcançando até 106 pontos a mais na média.

No que diz respeito ao rendimento familiar, percebe-se que o percentual de famílias que se situam nas camadas mais altas de renda (classes a e b) é maior na região Sudeste do que na Nordeste, enquanto nas camadas mais baixas, ocorre exatamente o inverso, com destaque para a classe e, em que cerca de 78% das famílias dos alunos do 3º ano do ensino médio que realizaram o exame do ENEM se encontram recebendo até 2 salários mínimos, em contraposição ao Sudeste, que possui 48% das famílias nestas condições de renda. A implicação deste fato no desempenho escolar se faz presente na demanda por uma educação de qualidade, visto que uma maior renda implica em maiores gastos com educação.

Outro fator que corrobora com um bom desempenho educacional e que pode explicar as diferenças presentes entre ambas as regiões é o grau de instrução dos pais, principalmente o nível educacional da mãe. Percebe que o percentual de mães que possuem ensino superior na região Nordeste é de 14.34%, enquanto 21% das mães dos alunos que realizaram o ENEM na região Sudeste o possuem. Este descompasso educacional fica mais evidente quando se analisa o percentual de mães que possuem somente o ensino fundamental. 51% das mães de estudantes nordestinos possuem apenas o nível fundamental de educação, valor este, maior que o do Sudeste, que é de cerca de 43%.

O acesso à tecnologia e à informação também é considerado um fator preponderante no desempenho estudantil, visto que os custos de transação para se obter conhecimento se reduzem consideravelmente. Portanto, o acesso à internet e à computadores possui papel importante na educação moderna. Percebe-se que os estudantes nordestinos estão muito aquém no que diz respeito ao acesso à tecnologias da informação em contraposição aos alunos do Sudeste. Apenas 55% dos candidatos ao ENEM que moram no Nordeste possuíam computadores e internet, em contraposição ao Sudeste, com um percentual de 82%.

Em relação ao tipo de escola frequentada pelo aluno (pública ou privada), é consenso de que a qualidade do ensino ofertada nas escolas públicas é bem inferior aos da privada na média, portanto, presume-se que quanto mais alunos estudando em escolas públicas, menores serão seus rendimentos escolares. Percebe-se que o percentual de alunos que estudam em escolas públicas no Nordeste é maior que no Sudeste (cerca de 78% na primeira e 74% na segunda), corroborando portanto, para um menor desempenho nas provas de matemática do ENEM.

Após a análise descritiva dos dados, estimou-se o modelo hierárquico compreendendo o nível socioeconômico dos estudantes e o nível da escola. Porém, antes da estimação, foram realizados alguns procedimentos de adequação do modelo.

Com o intuito de confirmar a validade do modelo, foi realizado um teste χ^2 , o LR *test*, que tem como hipótese nula o fato dos efeitos aleatórios serem iguais a zero. Pôde-se constatar a rejeição da hipótese nula a um nível de significância de 1%, justificando mais uma vez a utilização de um modelo com estrutura hierárquica.

Os modelos hierárquicos mais difundidos são os de intercepto aleatório e os de inclinação aleatória. O primeiro decompõe a variância das variáveis independentes entre os níveis e o intercepto é aleatório, enquanto o segundo testa os efeitos aleatórios dos coeficientes angulares. Para garantir que o modelo de intercepto aleatório fosse realmente o melhor modelo a ser estimado, foi realizado um teste de razão de máxima verossimilhança a fim de comparar o modelo de intercepto aleatório com um modelo de intercepto e inclinação aleatórios. Para tal, estimou-se um modelo em que o nível de renda da família é função do tipo de escola que o aluno frequenta. A hipótese nula do teste é de que os modelos não apresentam diferenças significativas contra a hipótese alternativa de que eles são distintos. Os resultados indicam a não rejeição da hipótese nula (p -valor= 0,7271), demonstrando que os modelos comparados não apresentam diferenças, podendo então optar pelo modelo de intercepto aleatório. Além disso, os componentes cruzados da estimação com inclinação aleatória se mostraram não significativos, corroborando ainda mais com a utilização do modelo de intercepto aleatório.

Partindo agora para a estimação do modelo de intercepto aleatório, primeiramente estima-se o modelo nulo, ou seja, sem nenhuma das variáveis explicativas, deixando apenas o intercepto. Estima-se o modelo nulo somente à título de comparação. Isto permite a análise da distribuição da variância total do desempenho em matemática pelos níveis de agrupamento sem os efeitos das variáveis explicativas. Pressupõe-se que ao se estimar o modelo nulo, a variância das notas explicada pelo agrupamento em tipos de escola seja maior, visto que ao se incluir variáveis explicativas no primeiro nível, a importância do segundo nível cai. A estimativa da variância entre as escolas foi 2,998.976, enquanto a variância entre alunos foi 6,517.646. Desta forma, foi possível a construção do coeficiente intra-classe, que descreve o quanto da variação total das notas em matemática se deve à diferença entre escolas. O coeficiente foi calculado através do quociente entre a variância do nível da escola e a variância total. O resultado, apresentado na tabela 2, nos mostra que 31.51% da variação total do desempenho escolar é explicada pela variação entre escolas.

Ao se estimar o modelo completo, agora inserindo as variáveis explicativas do primeiro e segundo níveis, percebe-se uma redução do coeficiente intra-classe para 0.1528. Ou seja, apenas 15.28% da variância total das notas dos alunos do nordeste é explicada pela diferença entre escolas públicas e privadas. Tal redução era de se esperar ao serem incluídas variáveis independentes, reduzindo assim o percentual da variância total explicada pelo tipo de escola, evidenciando a importância do nível do aluno para explicar o desempenho escolar.

Os resultados da estimação das variâncias e seus respectivos coeficientes intra-classe são apresentadas na tabela 3 a seguir:

Tabela 3: Análise da importância do tipo de escola no desempenho escolar no ENEM para os estudantes do Nordeste

	Modelo nulo	Modelo completo
Variância do primeiro nível	6,517.646	6,067.957
Variância do segundo nível	2,998.976	1,089.327
Variância total	9,516.622	7,157.284
Coeficiente intra-classe	0.3151	0.1528

Fonte: Resultados da pesquisa

A partir daí, pode-se verificar os efeitos que as variáveis explicativas têm sobre o desempenho do aluno. Para tal, são apresentados os resultados dos coeficientes estimados na tabela 4, bem como suas respectivas significâncias estatísticas e seus desvios padrões.

Tabela 4: Modelo dos determinantes do desempenho em matemática no ENEM para os estudantes do Nordeste em 2013.

Nota matemática	Coefficientes	Erro padrão	P-valor
Sexo	31.4423	0.2840	0.000
Cor	3.506071	0.3317	0.000
Membros domicílio	-0.8394	0.0939	0.000
Classe a	29.0322	1.5909	0.000
Classe b	19.9588	1.1869	0.000
Classe c	10.2539	0.6999	0.000
Classe e	-14.3507	0.4949	0.000
Trabalha	-7.6471	0.3675	0.000
Internet	3.6517	0.3916	0.000
Computador	8.9123	0.3851	0.000
Instrução sup. Mãe	23.1410	0.8351	0.000
Instrução méd. mãe	20.5695	0.7476	0.000
Instrução fund. Mãe	9.8401	0.6865	0.000
Instrução sup. Pai	21.0830	0.8031	0.000
Instrução méd. pai	15.8349	0.6282	0.000
Instrução fund. Pai	7.5191	0.5346	0.000
Escola pública	-55.4959	1.0513	0.000
Constante	496.1722	1.2842	0.000

Fonte: resultados da pesquisa

Percebe-se que todos os coeficientes estimados são estatisticamente significativos a um nível de 1% e os sinais corroboram exatamente com a literatura vigente. Podem-se fazer algumas inferências a respeito dos resultados apresentados. No que diz respeito à educação dos pais, percebe-se uma maior importância do nível de

instrução da mãe do que o do pai para todos os diferentes níveis (fundamental, médio e superior). Alunos com mães que possuem ensino superior completo tem um aumento de 23.1410 pontos na nota de matemática, que representa uma evolução em 4.66% da nota, enquanto aqueles que possuem pais com ensino superior têm um aumento ligeiramente menor, de 21.0830 pontos, ou 4.24%. Isto corrobora com o fato já confirmado pela literatura de que a mãe tem um papel central e fundamental na educação dos filhos, dedicando mais tempo aos estudos dos filhos do que os pais, além de pressioná-los a alcançarem melhores resultados e níveis mais altos de educação. Através das estatísticas descritivas percebe-se que a região Nordeste possui uma baixa escolaridade dos pais se comparada à região Sudeste. Apenas 14% das mães de alunos nordestinos possuem ensino superior, em contrapartida, o Sudeste possui 21%.

A variável renda também é fundamental quando se trata do desempenho do aluno. Percebe-se que famílias com uma renda familiar mensal que se enquadra na classe a (mais de 20 salários mínimos) têm um aumento de 29.0321 pontos da prova de matemática do ENEM, ou seja, 5.85% de aumento na média, enquanto aquelas famílias mais pobres, que se situam na classe e (até dois salários mínimos) tem sua nota reduzida em 14.3507 pontos, aproximadamente 3%. Isto mostra a importância da renda ao se demandar educação, visto que esta é capaz de cobrir os custos diretos e indiretos de uma educação de qualidade, proporcionando conseqüentemente, melhores resultados. Além disso, um maior nível de renda pode proporcionar um maior conforto, possibilitando ao aluno uma maior dedicação de seu tempo escasso aos estudos.

Fatores como a presença de computadores e internet no domicílio do aluno também são fundamentais para o desempenho escolar, visto que estas ferramentas reduzem consideravelmente o custo ao acesso à informação. A presença de um microcomputador e internet no domicílio aumenta a nota do estudante no ENEM em 12.564 pontos, o que equivale a 2.5% de aumento na média final.

Variáveis de controle como o sexo e a cor também se mostraram estatisticamente significativos e com efeitos positivos sobre o desempenho escolar. Alunos do sexo masculino e brancos tem um aumento na nota de matemática de 34.9484 pontos, o equivalente a 7% da média final. Como se pôde identificar através das estatísticas descritivas, a proporção de alunos brancos e não brancos na região Nordeste é desigual, sendo a maioria pardos e negros (78%). Este fato se torna ainda mais relevante quando analisamos a proporção de brancos e não brancos que compõem a classe E de renda. 73.66% dos alunos que possuem rendimento familiar de até dois salários mínimos são não brancos, demonstrando assim que o fator racial está muito interligado a má distribuição de renda.

O trabalho na adolescência, como era de se esperar, afetou negativamente o desempenho dos estudantes. Para aqueles que trabalham ou já trabalharam, a nota final em matemática no exame do ENEM tem uma redução de 7.6471 pontos, cerca de 1.5% da média final, demonstrando que a alocação do tempo do estudante no trabalho em detrimento ao estudo é perversa quando se trata do rendimento escolar, já que a dedicação aos estudos fica prejudicada pela necessidade de trabalhar.

Por fim, o fator de maior relevância no desempenho dos alunos do 3º ano do ensino médio da região Nordeste foi o fato da escola ser pública ou privada. Verifica-se, que para as escolas públicas, há uma redução de 55.4959 pontos na média final do aluno, um total de 11.2% da média total. Tal resultado pode ser justificado graças ao fato de a maioria dos alunos da rede pública nordestina ser negra, ter rendimento familiar de até dois salários mínimos e baixa escolaridade dos pais. Cerca de 91% dos alunos inseridos na rede pública de ensino são negros e tem rendimento familiar de até

dois salários mínimos, podendo justificar portanto, o menor rendimento das escolas públicas nordestinas.

V. CONCLUSÕES

O sistema educacional brasileiro apresenta heterogeneidades evidentes, havendo a necessidade, portanto, de realizar trabalhos investigativos desagregados em níveis. Foi exatamente isto que este artigo propôs, analisar o desempenho escolar na prova de matemática do ENEM dos alunos do 3º ano do ensino médio da região Nordeste tanto a nível do aluno quanto da escola. Para tal, utilizou-se de modelos hierárquicos em dois níveis, o primeiro contendo as características socioeconômicas do aluno e o segundo referente ao tipo de escola (pública ou privada).

Através do coeficiente intra-classe, foi possível determinar o tamanho do impacto do tipo de escola no desempenho escolar. Encontrou-se que 15.28% da variância total das notas de matemática podem ser explicadas pela diferenciação entre escolas.

No que diz respeito às variáveis socioeconômicas do estudante situadas no primeiro nível, verificou-se que todas as variáveis selecionadas foram significativas a 1% de nível de significância. Fatores como a maior escolaridade da mãe, maior renda, presença de computador e internet e se o estudante é branco e do sexo masculino tiveram impactos positivos no desempenho na prova de matemática do ENEM. Em contrapartida, variáveis como o trabalho na adolescência, o nível de renda baixo, o número de membros no domicílio e o fato de estudar em escola pública colaboraram para impactar negativamente na nota média do aluno.

Pode-se perceber, portanto, a importância da condição socioeconômica do aluno no desempenho escolar, que pode muito bem compensar os efeitos negativos que o tipo de escola pode gerar em termos de proficiência em matemática.

VI. REFERÊNCIAS BIBLIOGRÁFICAS

ALBERNAZ, Â., FERREIRA, F. H., & FANCO, C. (2002). Qualidade e equidade na educação fundamental brasileira. **Textos para discussão n° 455, PUC-Rio.**

ANDRADE, J. M., & LAROS, J. A. (2007). Fatores associados ao desempenho escolar: estudo multinível com dados do SAEB/2001. *Psicologia: Teoria e Pesquisa*, **23(1)**, 033-042.

BARROS, R. P. D., MENDONÇA, R., SANTOS, D. D. D., & QUINTAES, G. (2001). Determinantes do desempenho educacional no Brasil. **Textos para discussão n° 834, IPEA.**

BECKER, G.S. Investment in Human Capital: A Theoretical Analysis. **The Journal of Political Economy**. Vol. 70, No. 5, Oct. 1962.

BERQUÓ, E., CAVENAGHI, S. (2004). Mapeamento sócio-econômico e demográfico dos regimes de fecundidade no Brasil e sua variação entre 1991 e 2000. **Encontro Nacional de Estudos Populacionais**, *14*, 1-18.

BIONDI, R. L.; FELÍCIO, F. Atributos escolares e o desempenho dos estudantes: uma análise em painel dos dados do Saeb. **Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2007.**

CRUZ, C. C. M. S. (2010). Modelos multi-nível: fundamentos e aplicações (**Doctoral dissertation, Universidade Aberta**).

HANUSHEK, E. A. (1989). The impact of differential expenditures on school performance. **Educational researcher**, 18(4), 45-62.

HOX, J. (2010). Multilevel analysis: Techniques and applications. **Routledge**.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados do Enem 2013**. Brasília: Inep, 2015. Disponível em: <<http://portal.inep.gov.br/basicalevantamentos- acessar>>. Acesso em: 30 mai. 2015.

JESUS, G. R. D., & LAROS, J. A. (2004). Eficácia escolar: regressão multinível com dados de avaliação em larga escala. **Avaliação Psicológica**, 3(2), 93-106.

LAM, D., DURYEA, S. (1999). Effects of schooling on fertility, labor supply, and investments in children, with evidence from Brazil. **Journal of Human Resources**, 160-192.

MENEZES-FILHO, N. Os Determinantes do Desempenho Escolar do Brasil. **Instituto Futuro Brasil, Ibmecc-SP e FEA-USP**.

MINCER, J. Investment in Human Capital and Personal Income Distribution. **Journal of Political Economy**. Vol.66, No.4 (Aug., 1958), pp.281-302

RESENDE, M. & WYLLIE, E. Retornos para Educação no Brasil: Evidências Empíricas Adicionais. **Revista Economia Aplicada**, Vol. 10, No. 3, Ribeirão Preto, July/Sept. 2006.

RIANI, J. D. L. R., & RIOS-NETO, E. L. G. (2008). Background familiar versus perfil escolar do município: qual possui maior impacto no resultado educacional dos alunos brasileiros. **Revista Brasileira de Estudos de População**, 25(2), 251-269.

SACHSIDA, A. ET. AL. Um Estudo Sobre Retorno em Escolaridade no Brasil. **Revista Brasileira de Economia**, Vol. 58, No. 2, Rio de Janeiro, Apr/June 2004.

SÁTYRO, N. G. D.; SOARES, D. S. **O impacto de infra-estrutura escolar na taxa de distorção idade-série das escolas brasileiras de ensino fundamental: 1998 a 2005**. IPEA, texto para discussão | 1338 | mai. 2008.

SCHULTZ, T.W. Investment in Human Capital. **The American Economic Review**. Vol. 51, No 1 (Mar.,1961),pp. 1-17.

SOARES, J.F. O Efeito da Escola no Desempenho Cognitivo dos Seus Alunos. **REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación 2004**, Vol. 2, No. 2.

UEDA, E.M. & HOFFMANN, R. (2002). Estimando o Retorno em Educação no Brasil. **Economia Aplicada**, 6(2): 209-238.